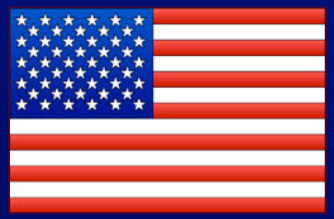
From Image-Space To Blob-Space:

the processing pipeline of FMRI data

Ziad S Saad, PhD

SSCC / NIMH & NINDS / NIH / DHHS / USA / EARTH



FMRI? It's easy!

Single subject time series

Expriment Design

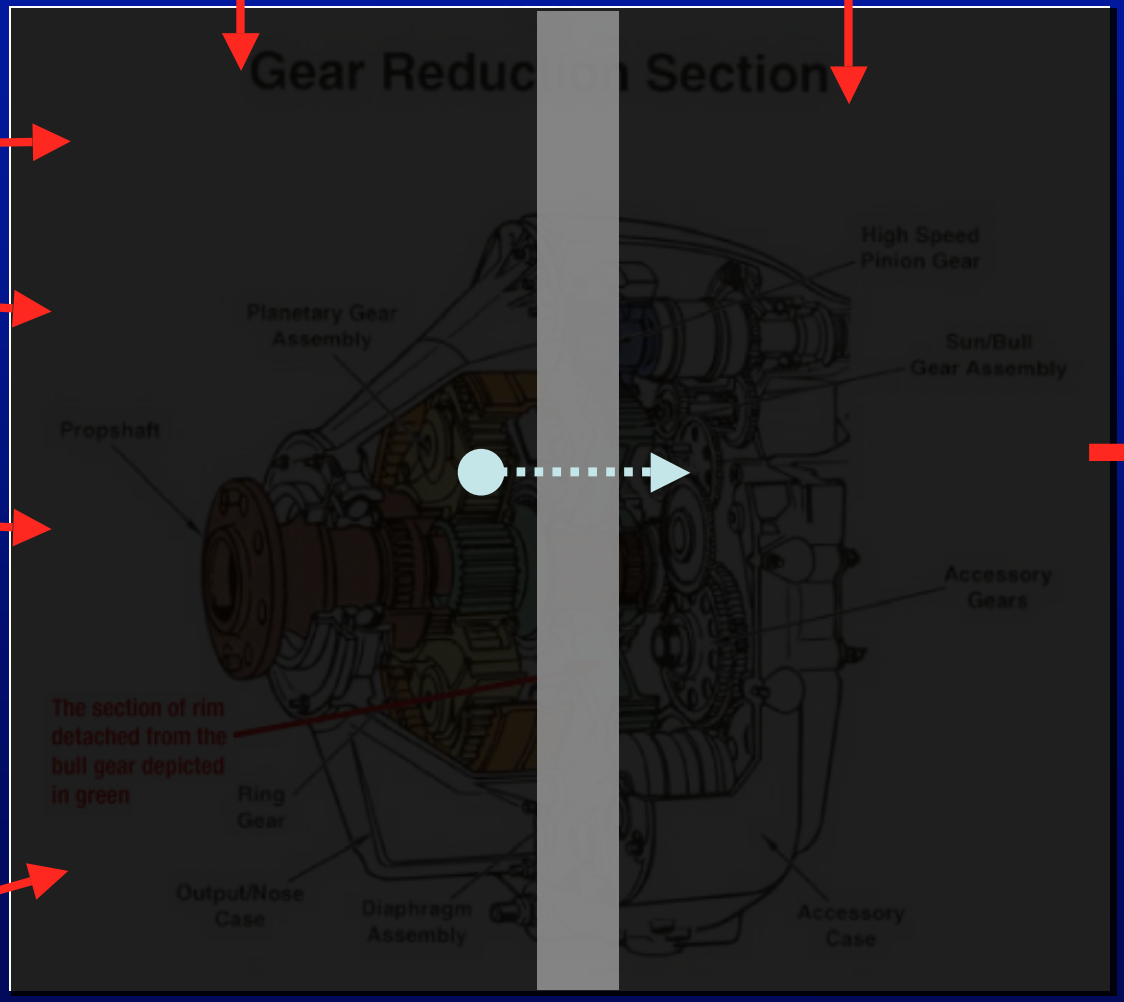
Contrast of interest

S.1

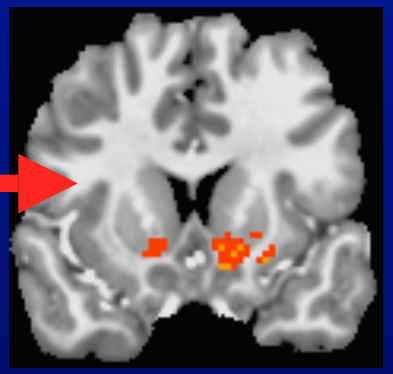
S.2

S.3

S.N



Group Results



FMRI? It's easy!

Single subject time series

Expriment Design

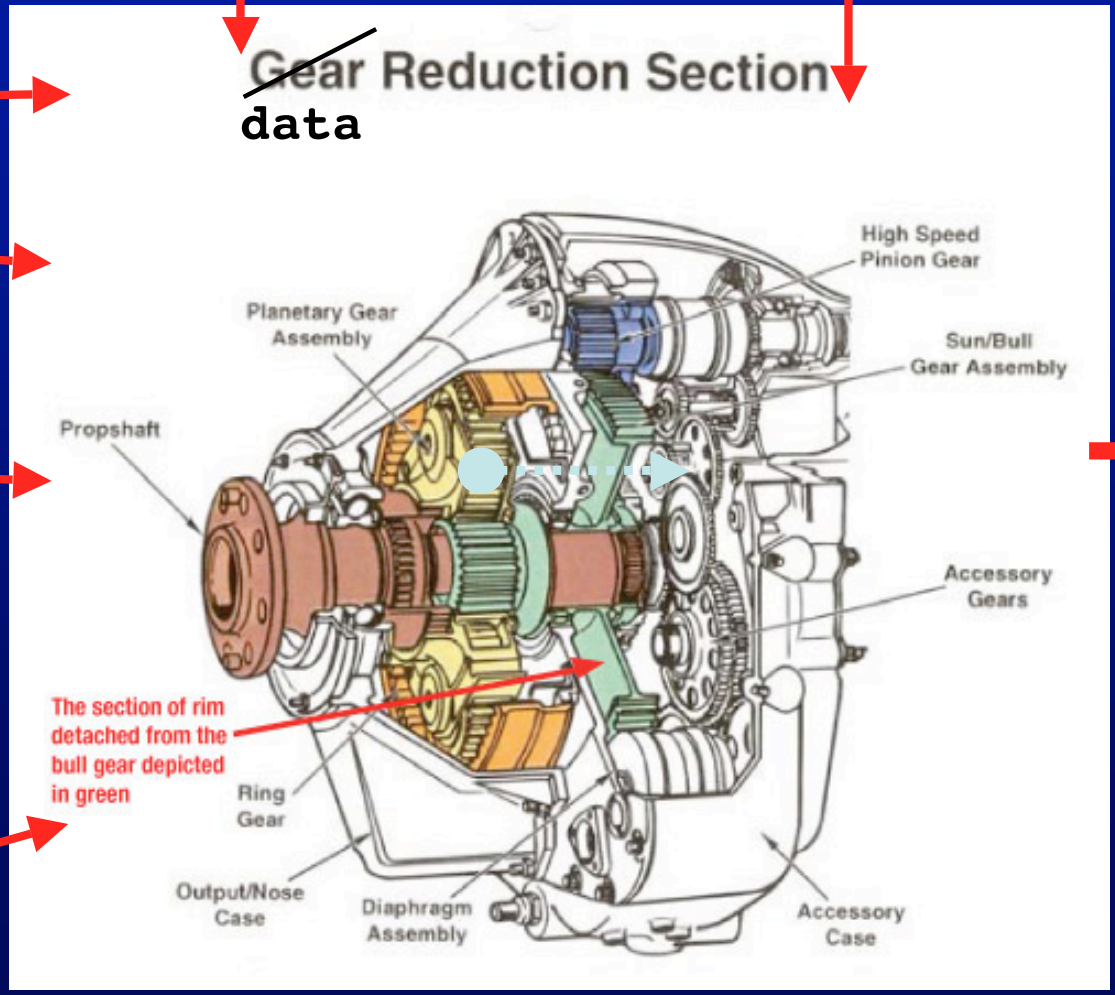
Contrast of interest

S.1

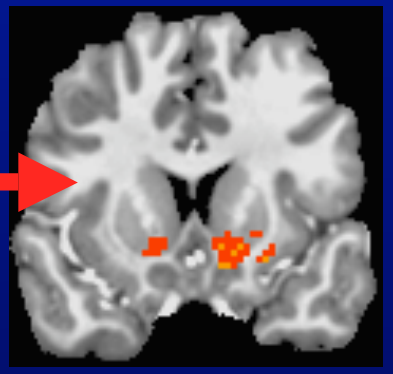
S.2

S.3

S.N



Group Results



Z.S.S 8-09

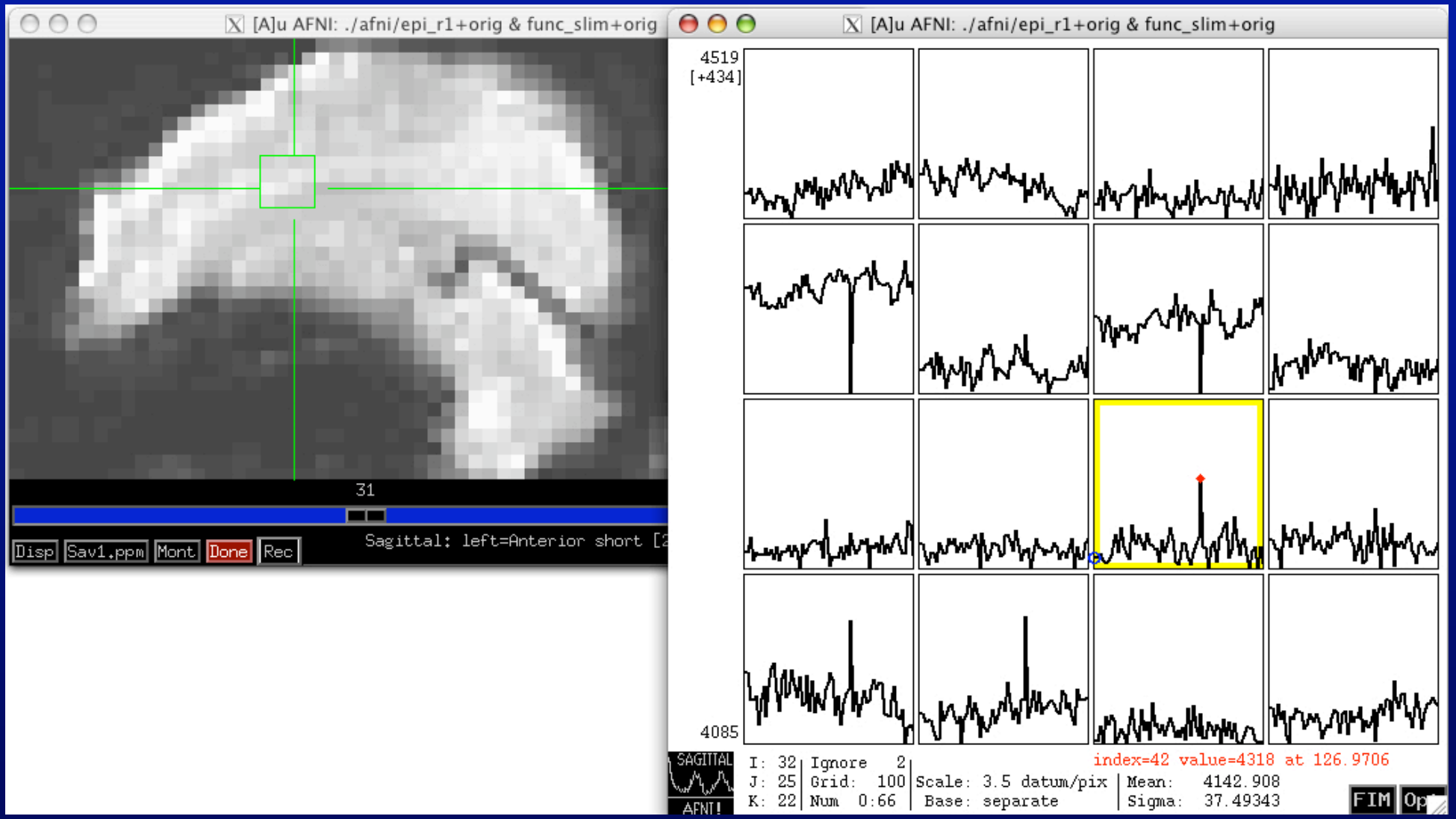
Stage 1- Single Subject Analysis

- **`afni_proc.py`** generates a Unix shell script to run a standard sequence of tools on an individual subject's time series datasets.
  - Despiking
  - RETROICOR-izing
  - Time shifting
  - **Volume registration**
  - **Blurring**
  - Mask generation [not applied at individual subject level]
  - EPI Scaling
  - **Regression analysis**
  - Spatial normalization
- Output datasets are ready for group-level analyses
- All processing blocks are optional and customizable
- Users are very encouraged to look at intermediary results
  - Data checking sequence automation

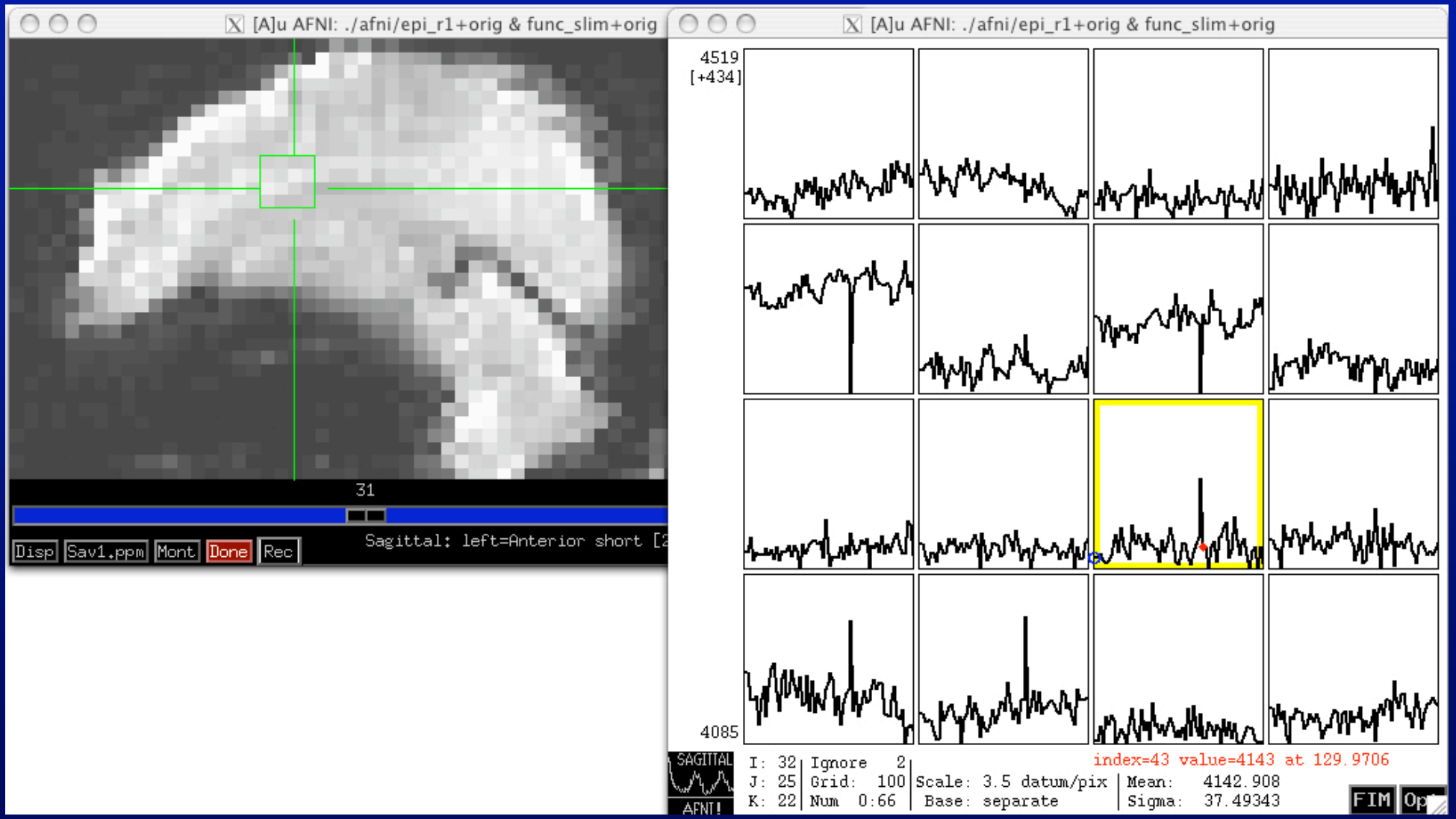
Spikes and other spiky things

- Not all spikes are created equal
  - Motion, usually OK
  - Hardware spikes need to be dealt with
  - Weirder artifacts appear with fancier equipment
- Spikes could get propagated with time series filtering, such as slice timing correction.
  - Reduce them before such operations
- Look at them before deciding what to do next
  - If they are due to motion, then motion regressors would absorb them
  - Could add a regressor for a spike, or just censor time point in latter analysis

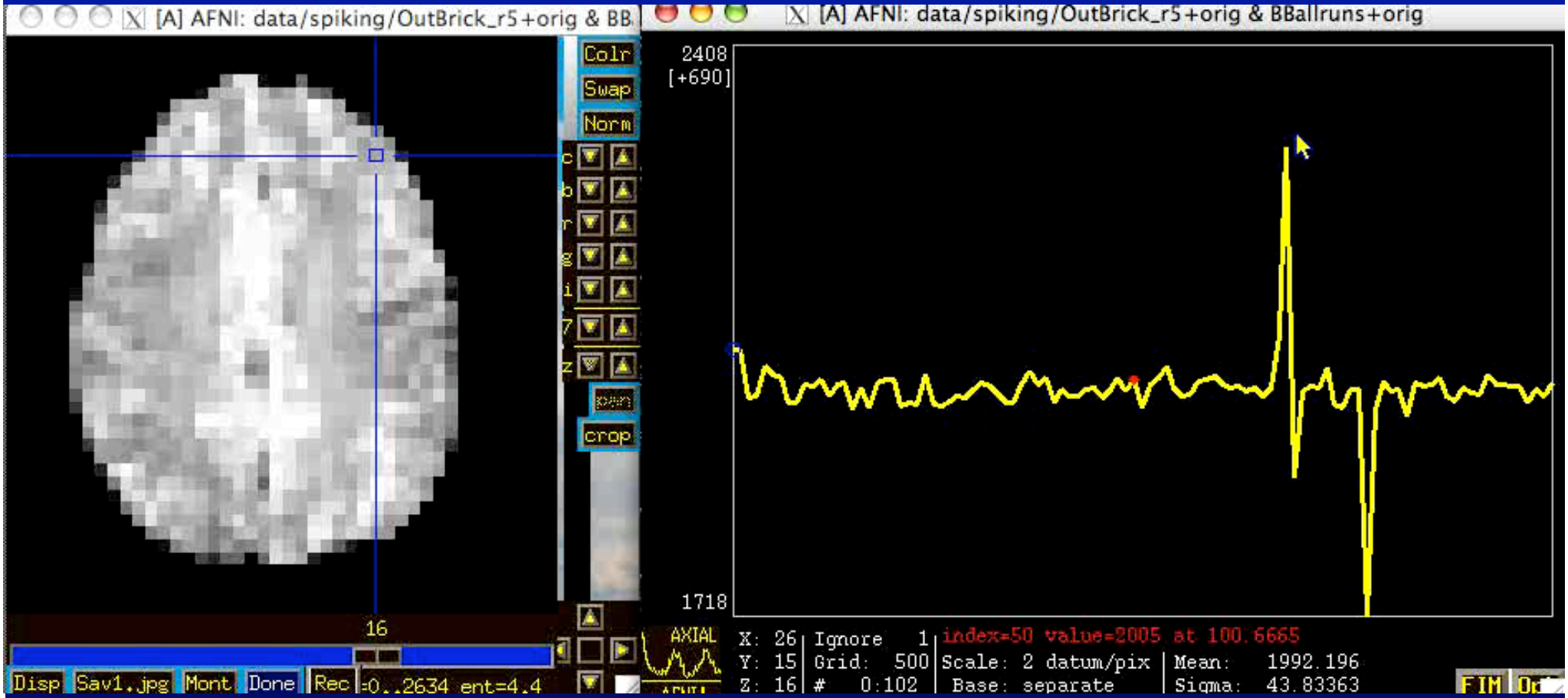
Movement Spikes



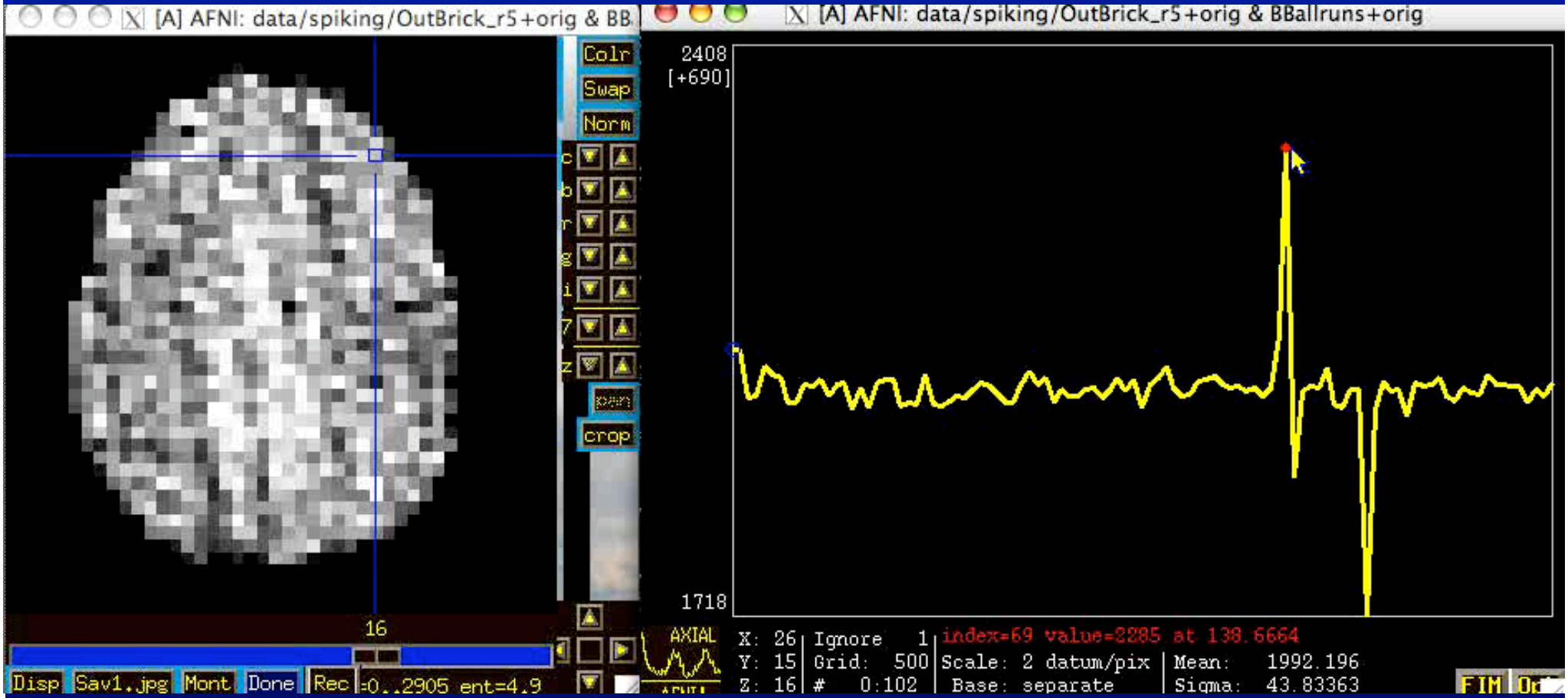
Movement Spikes



Hardware Spike

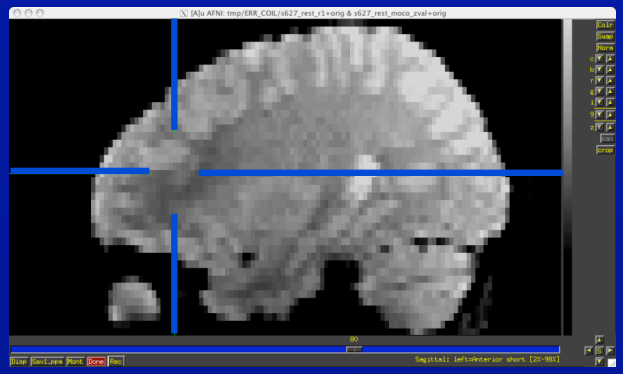


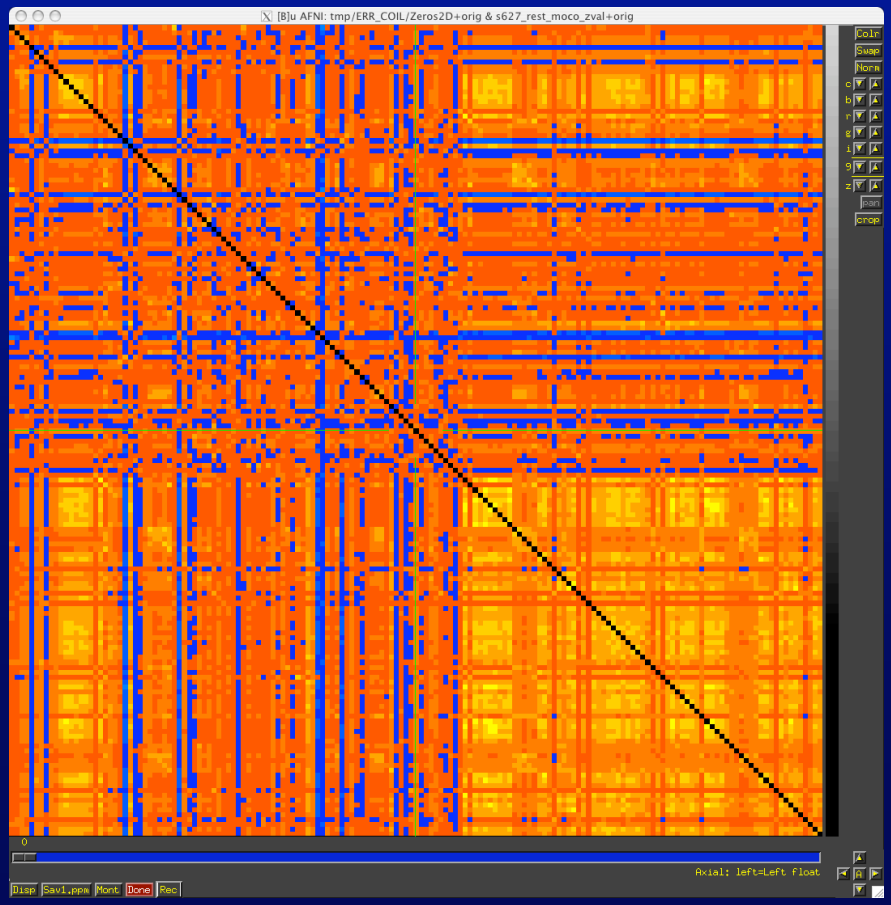
Hardware Spike

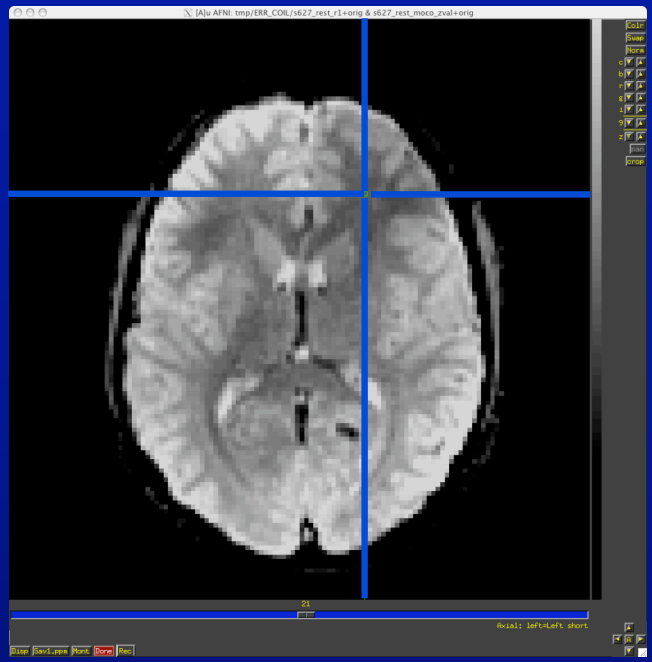


- Spikes caused by loose gradient coil connection

Weirder Spikes

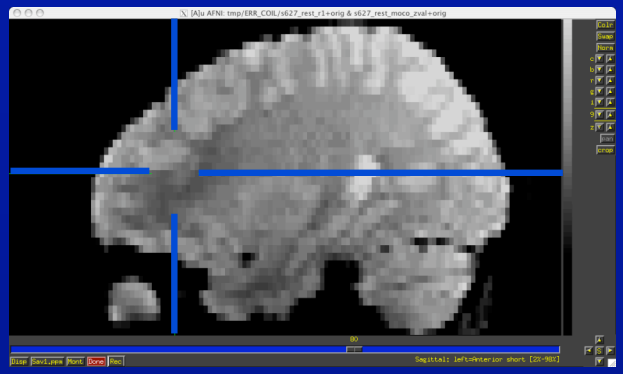


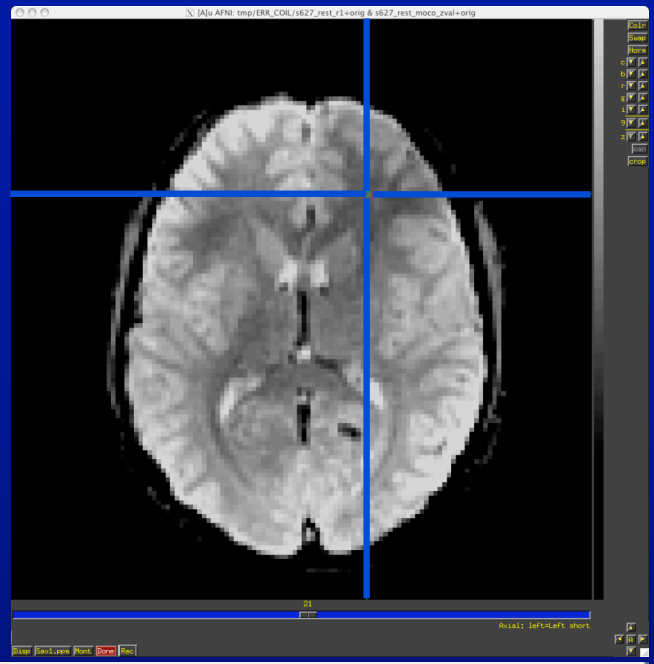


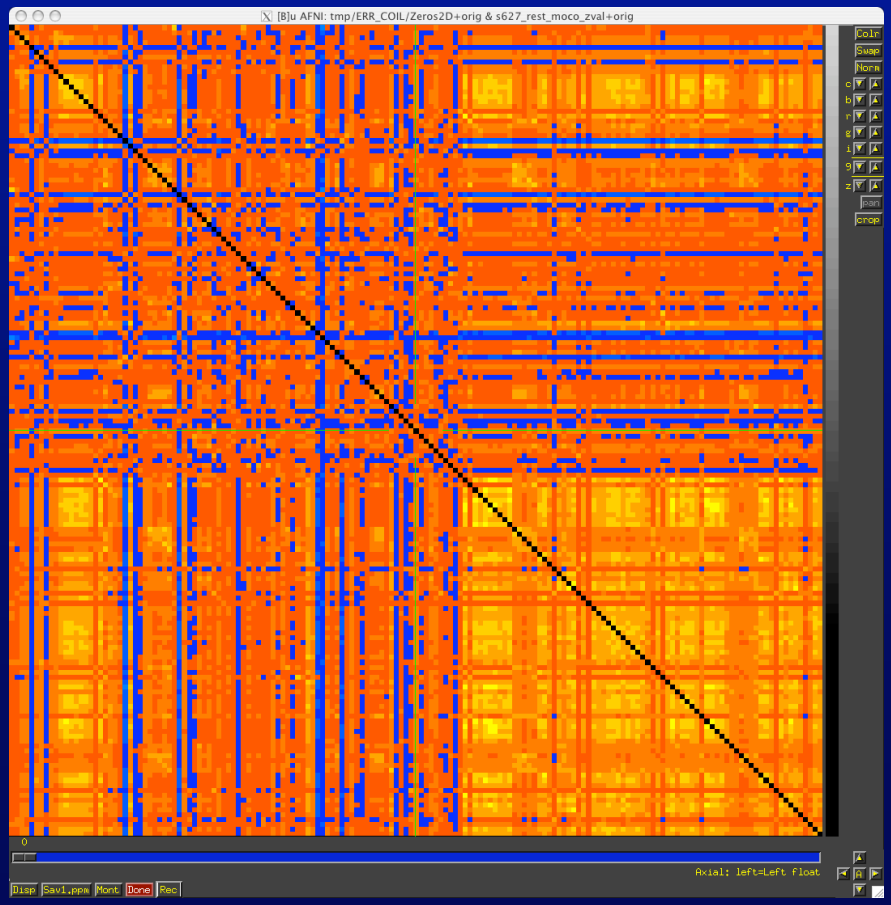


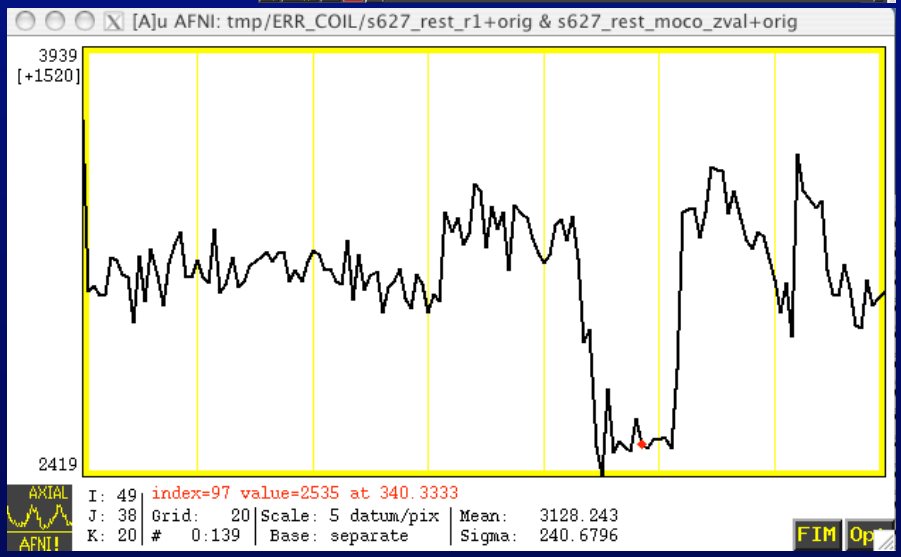
Z.S.S 8-09

Weirder Spikes

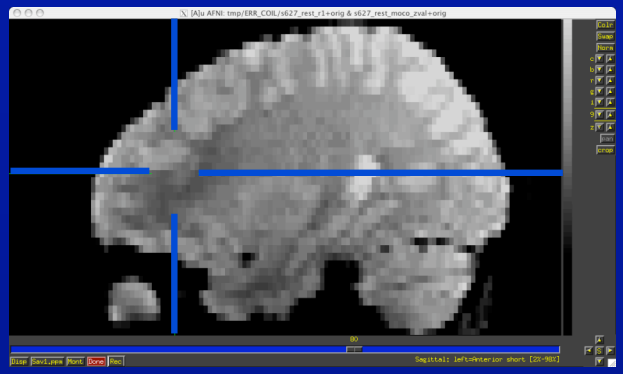


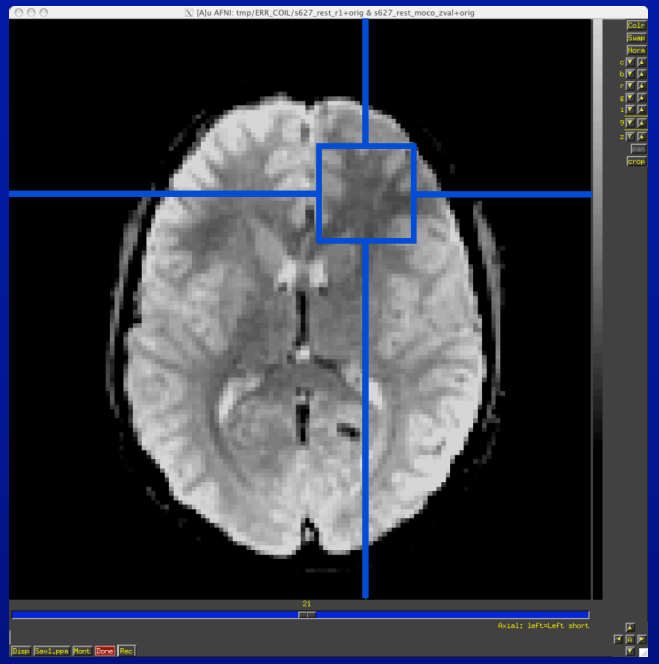


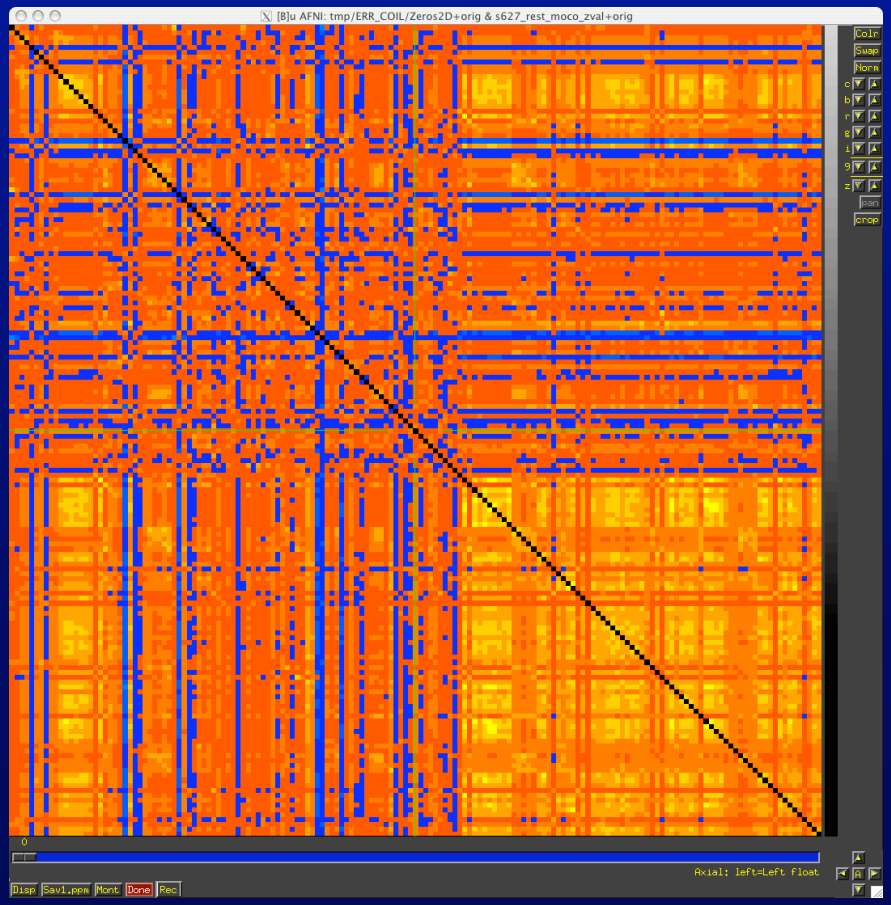


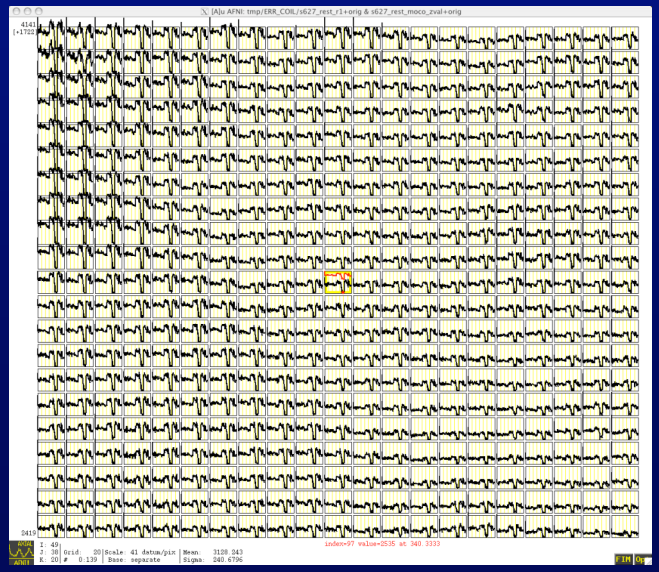


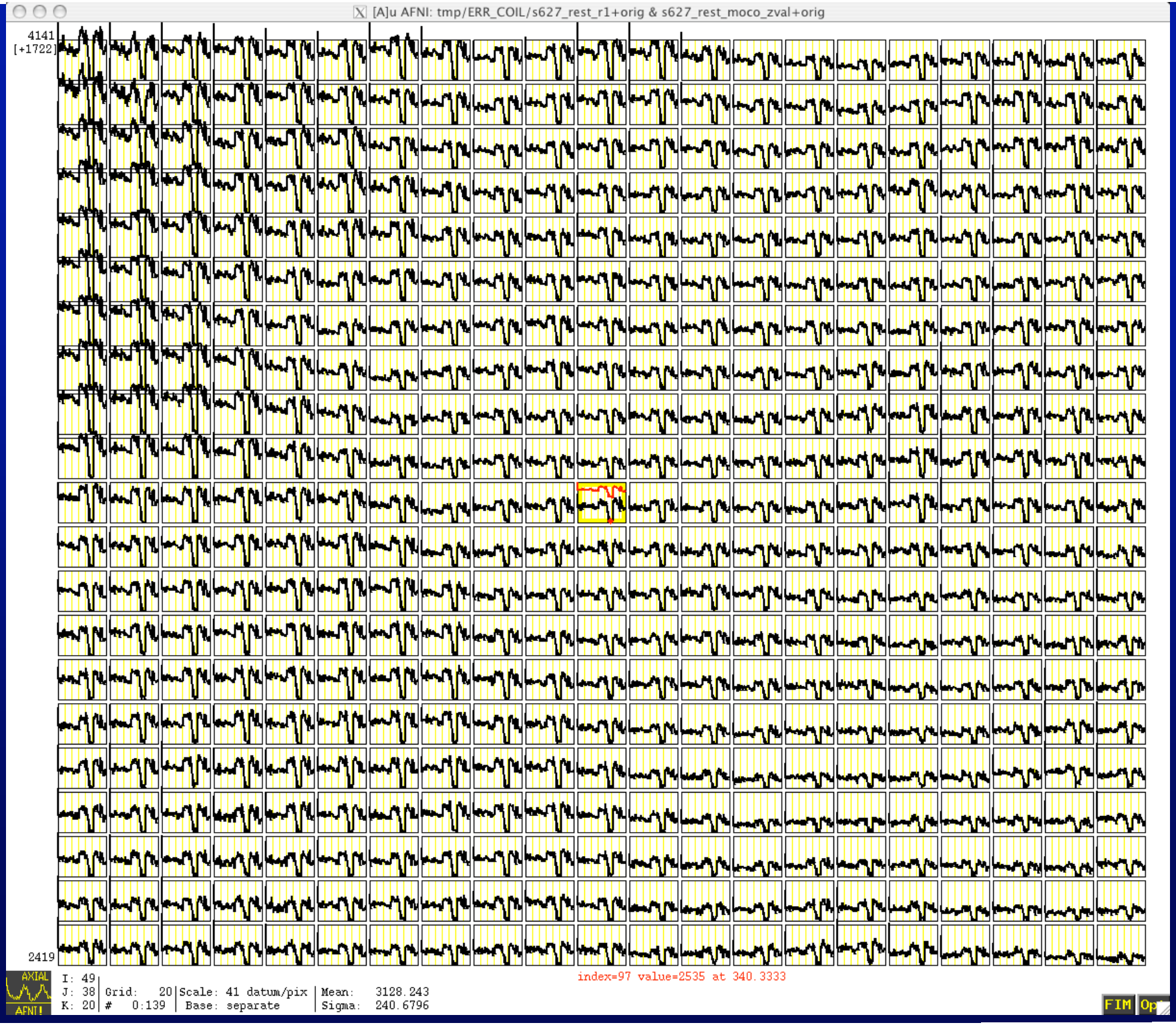
Weirder Spikes









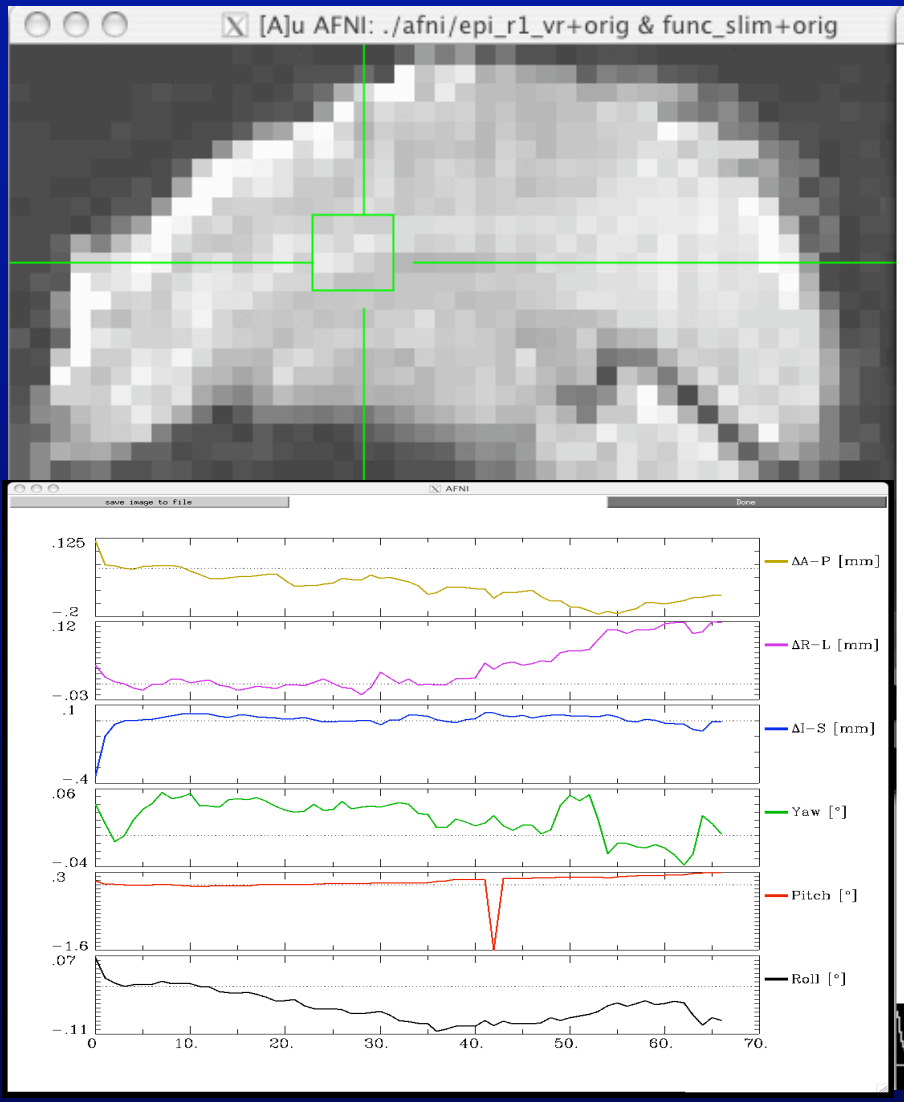

[A]u AFNI: tmp/ERR_COIL/s627_rest_r1+orig & s627_rest_moco_zval+orig
4141
[+1722]
2419
index=97 value=2535 at 340.3333
AXIAL
AFNI
I: 49
J: 38
K: 20
Grid: 20
0:139
Scale: 41 datum/pix
Base: separate
Mean: 3128.243
Sigma: 240.6796
FIM
Op

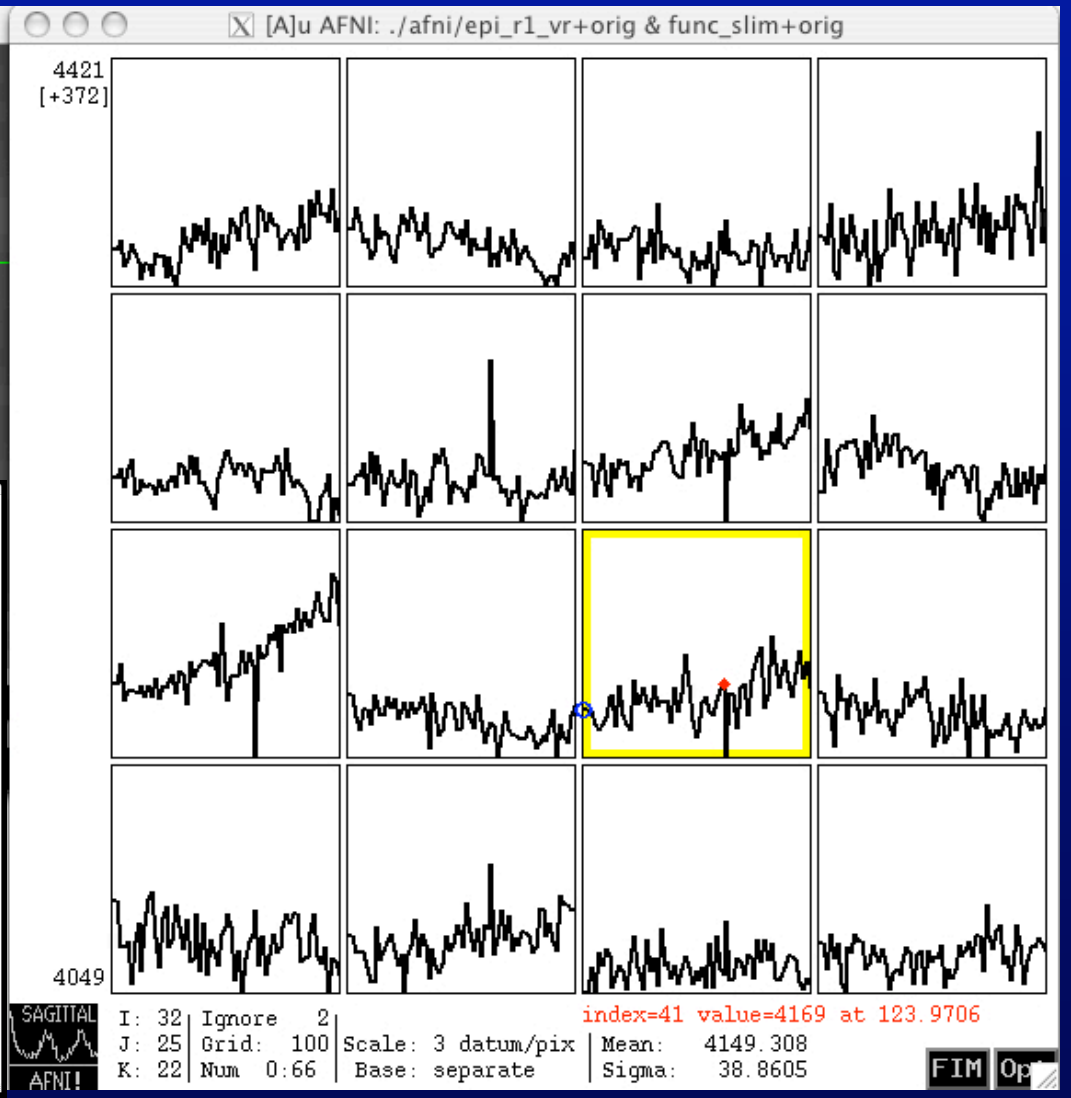
Motion Correction

- Within-modality: T2* to T2* or T1 to T1
  - Least squares cost functional is simple and robust
  - For EPI time series, rigid body (6 parameters) is typically used.
- Cross modality registration T1 to T2* for example
  - A variety of joint histogram based cost functionals
    - Elegant and general purpose.
    - But they can reach lowest cost at bad alignment
  - We propose the use of Local Pearson Correlation for an EPI to T1 cost functional

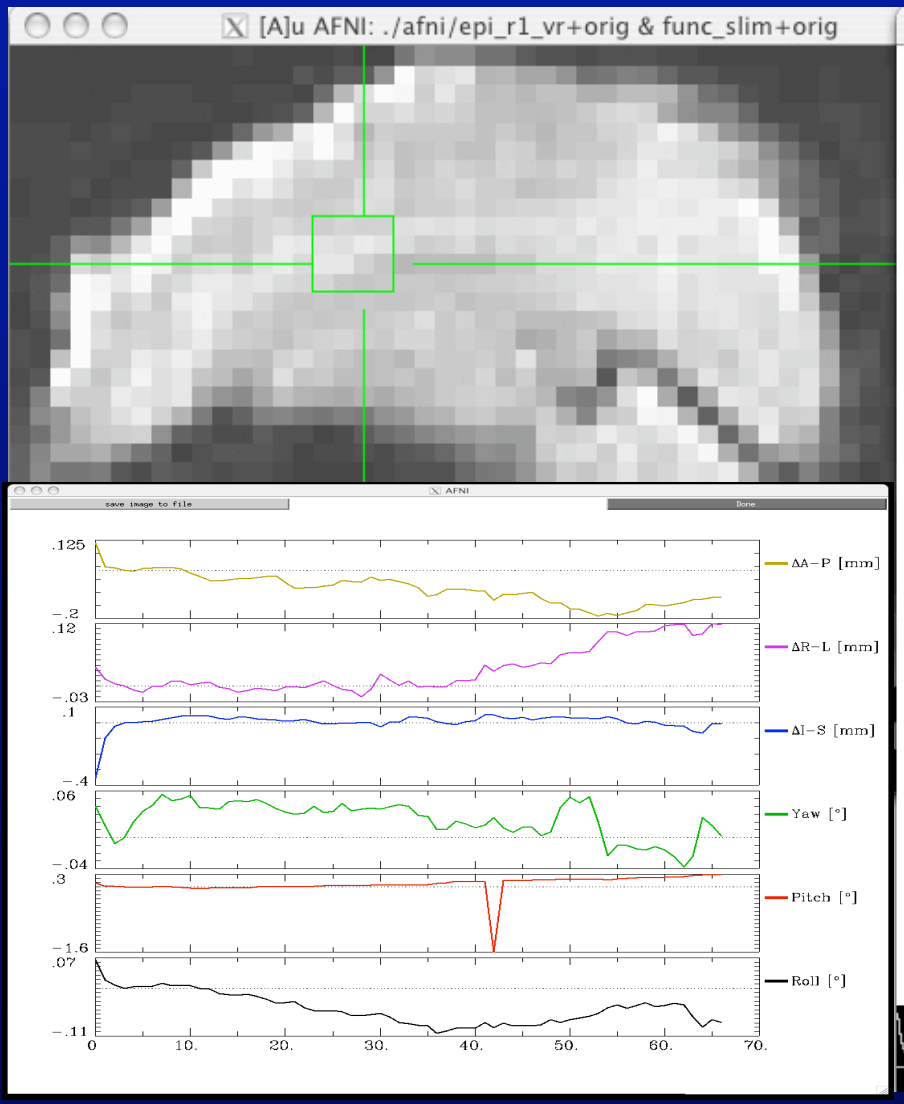
Movement Corrected

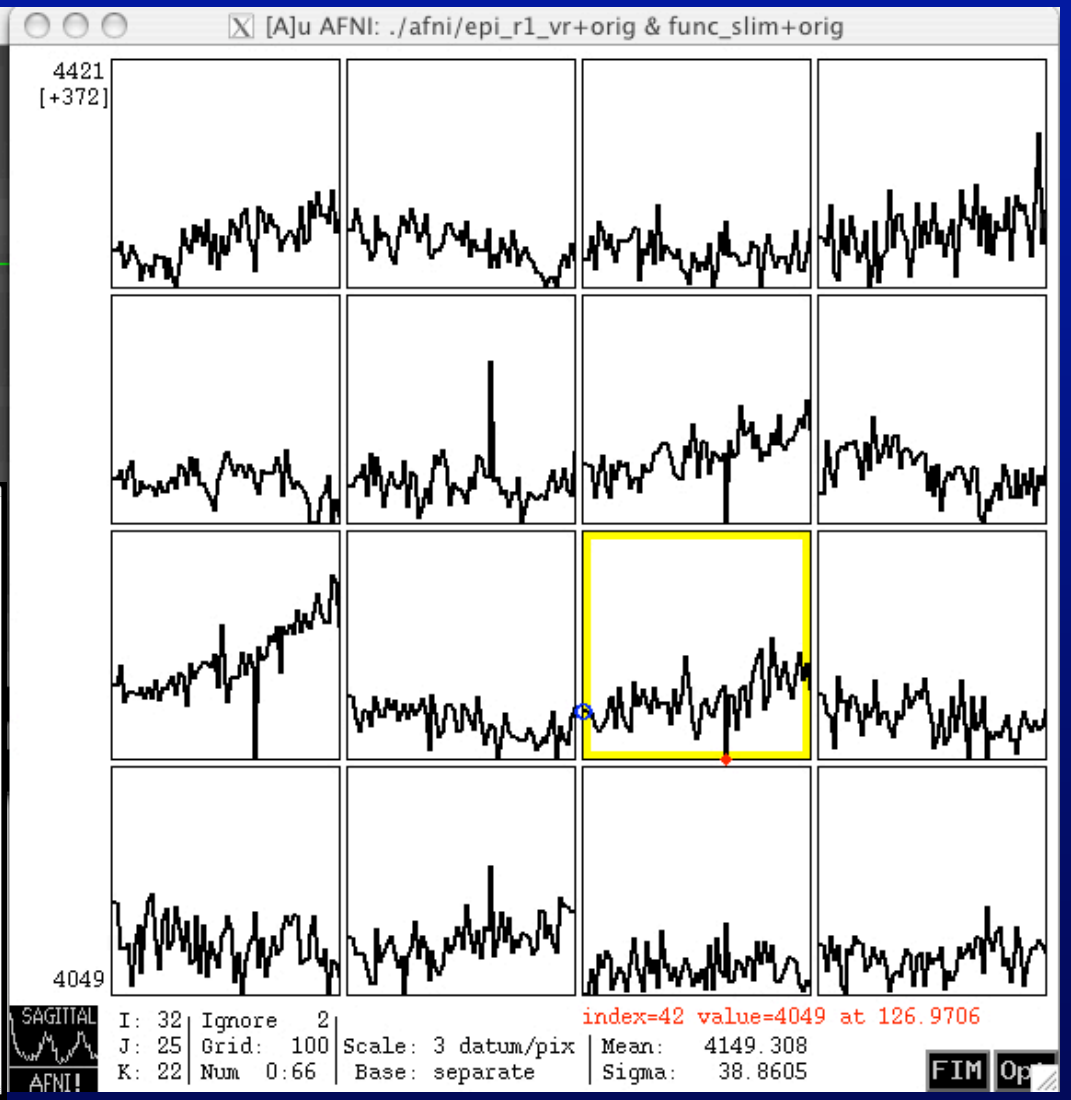
spikes remain





Movement Corrected
spikes remain





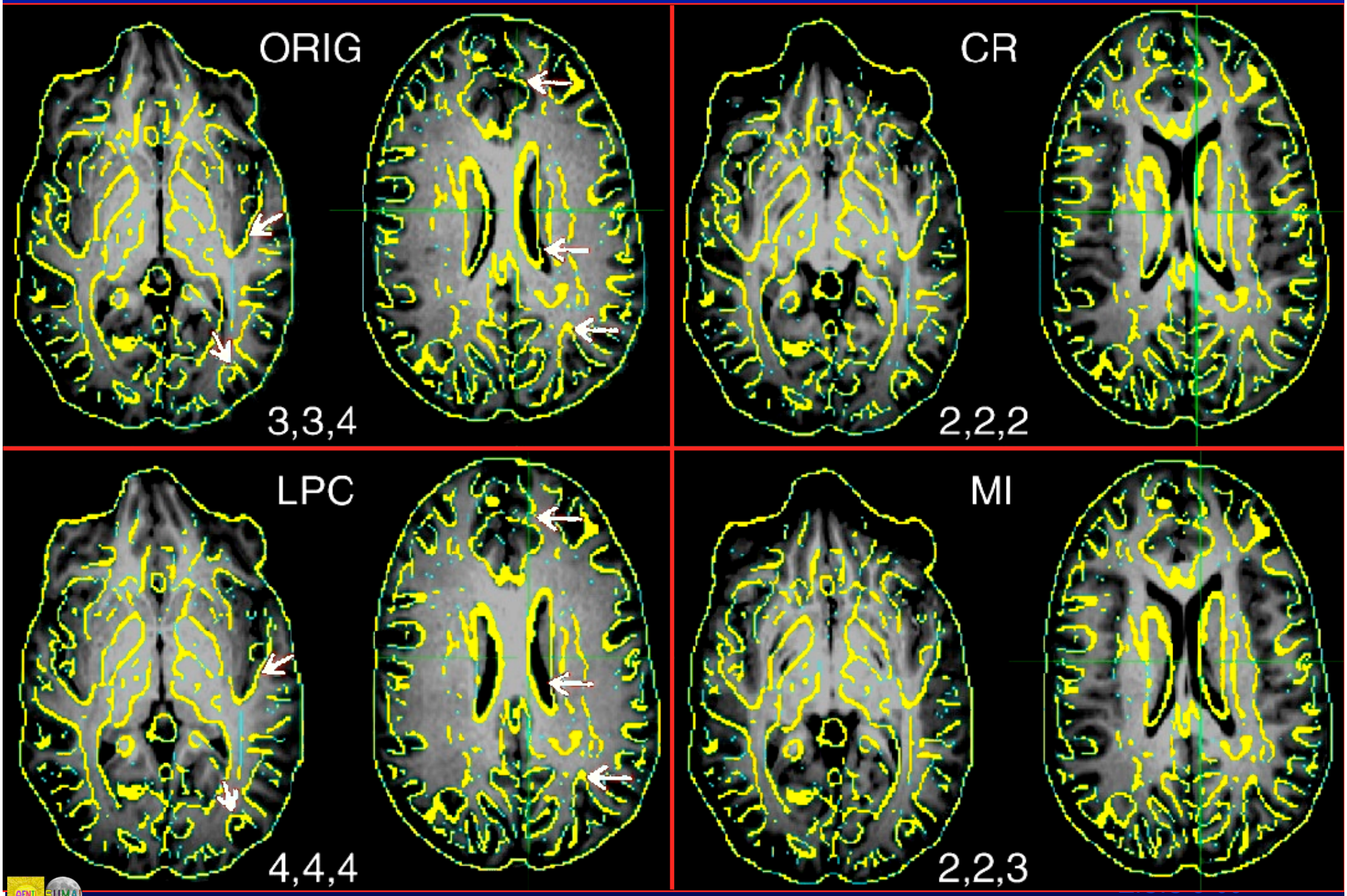
Motion Correction

- Within-modality: T2* to T2* or T1 to T1
  - Least squares cost functional is simple and robust
  - For EPI time series, rigid body (6 parameters) is typically used.
- Cross modality registration T1 to T2* for example
  - A variety of joint histogram based cost functionals
    - Elegant and general purpose.
    - But they can reach lowest cost at bad alignment
  - We propose the use of Local Pearson Correlation for an EPI to T1 cost functional

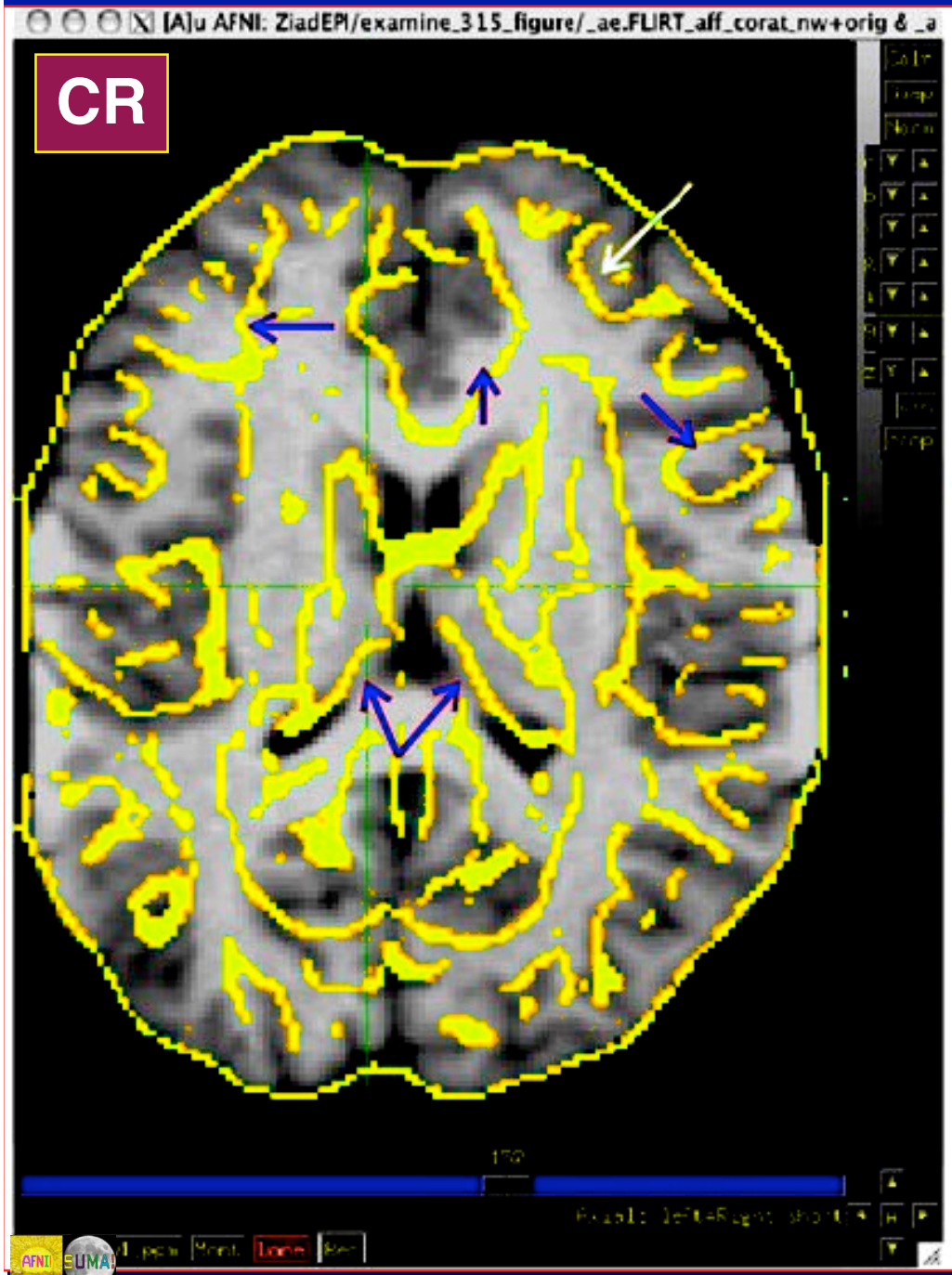
T2* ⇔ T1

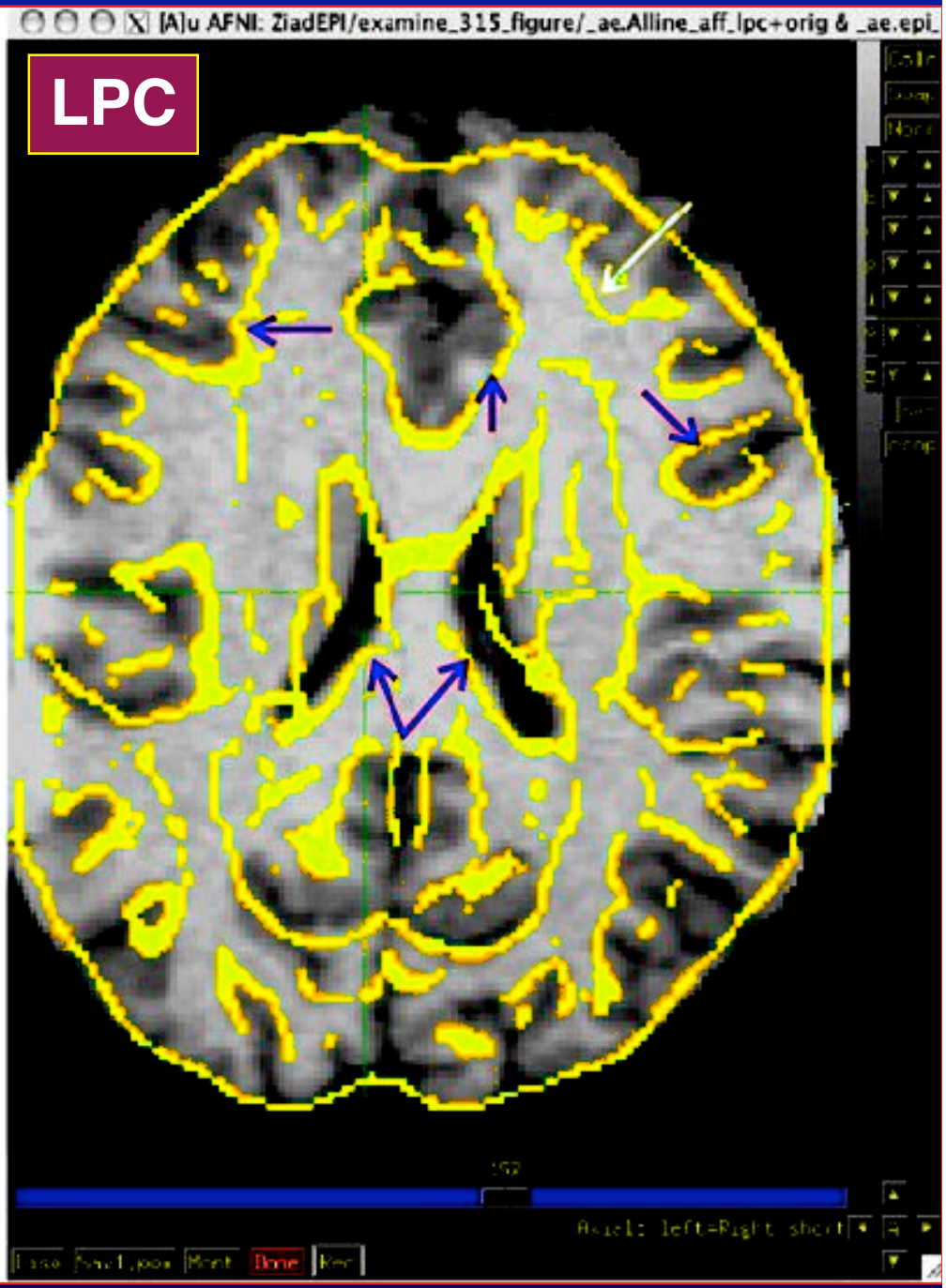
- Alignment of EPI T_2*-weighted volumes to structural T_1-weighted volumes (e.g., MPRAGE *or* SPGR) using generic inter-modality registration metrics (**MI**, **CR**) fails **10+%** of the time
  - Brain outlines may match, but internal structures can be **10+** mm away
  - Precise alignment needed for: cortical surface based analyses, use of anatomically defined ROIs, pre-surgical planning, …
- Local Pearson Correlation (**LPC**) metric (`3dAllineate`):
  - Compute correlation coefficients between EPI *and* structural volume ***locally*** over a collection of neighborhoods that cover the brain
  - Average this collection of correlations, weighted towards CSF regions (high intensity in EPI, low intensity in anatomical)
  - Optimization: adjust alignment until **LPC** is as ***negative*** as possible
  - Variant: |**LPC**| has been used to register 7T and 3T MPRAGEs
- More robust than **M**utual **I**nformation *or* **C**orrelation **R**atio
- **But**: user should ***always*** check image alignments visually***!!***
  - Edge enhanced image overlays are useful for this purpose

Results: EPI Edges Atop Anatomical Slices



Results: EPI Edges Atop Anatomical Slices



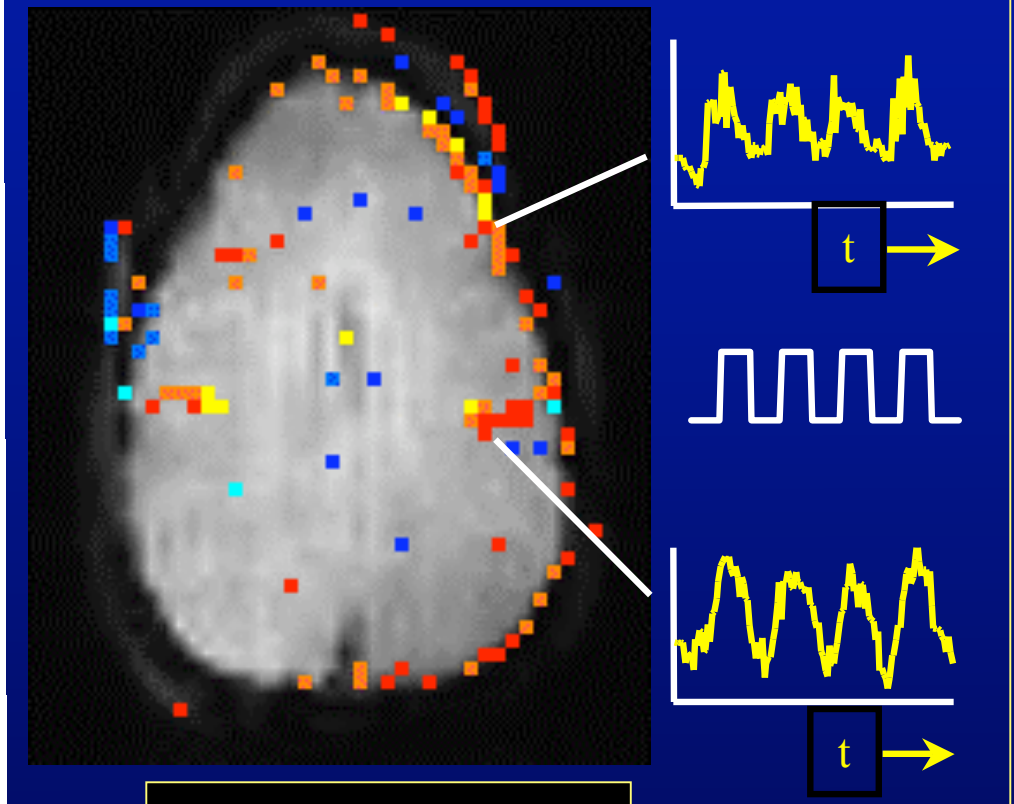


Stimulus Correlated Movement

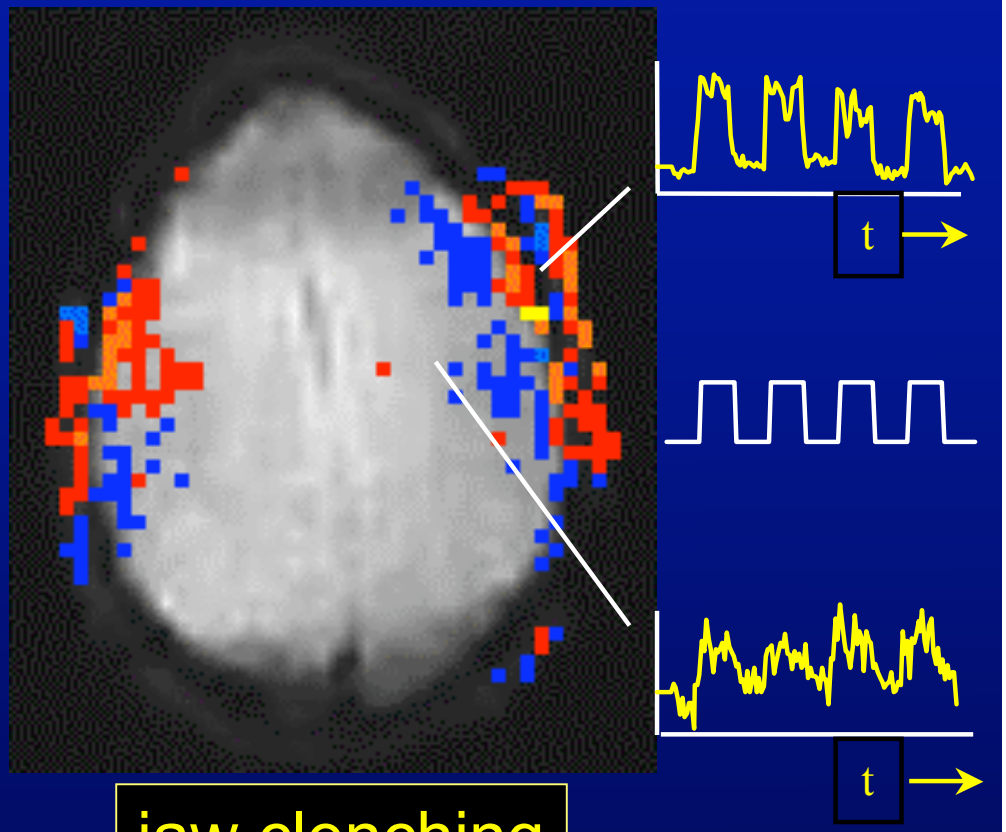
- By accident
  - Stimulus induced
    - Could confound results
      - Can happen in subtle ways as tensing up shoulders or changing breathing depth
      - Warning sign is stimulus-correlated signals on edge of brain
    - Careful consideration of stimulus timing can reduce this problem
  - Uncorrelated with Stimulus
    - Adds variance to data, resulting in less power
- By design
  - Speech production, swallowing, etc.

"Activation" Artifacts

R.M. Birn, et al. Human Brain Mapping 7(2), 106-114, 1999



overt speaking



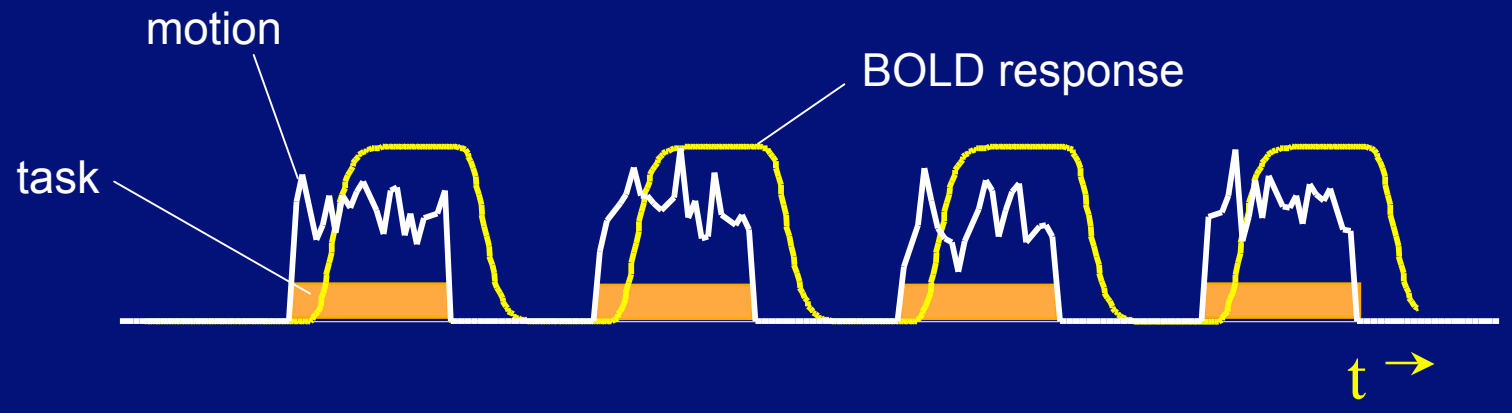
jaw clenching

- Non-BOLD signal changes correlated with task timing

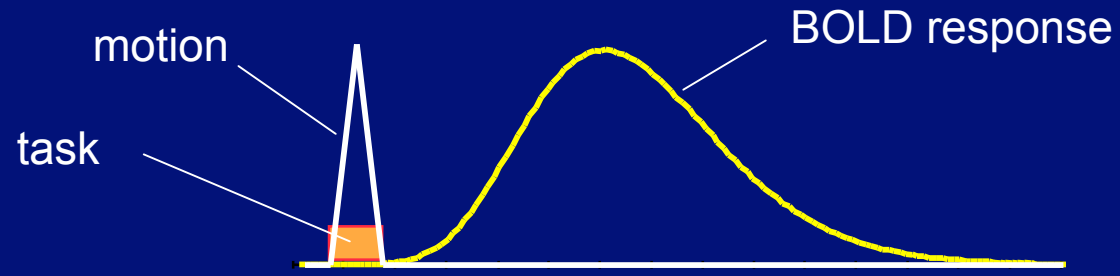
Slide courtesy of R. Birn

Motion Effect is Immediate
BOLD Effect is Delayed

Blocked Design

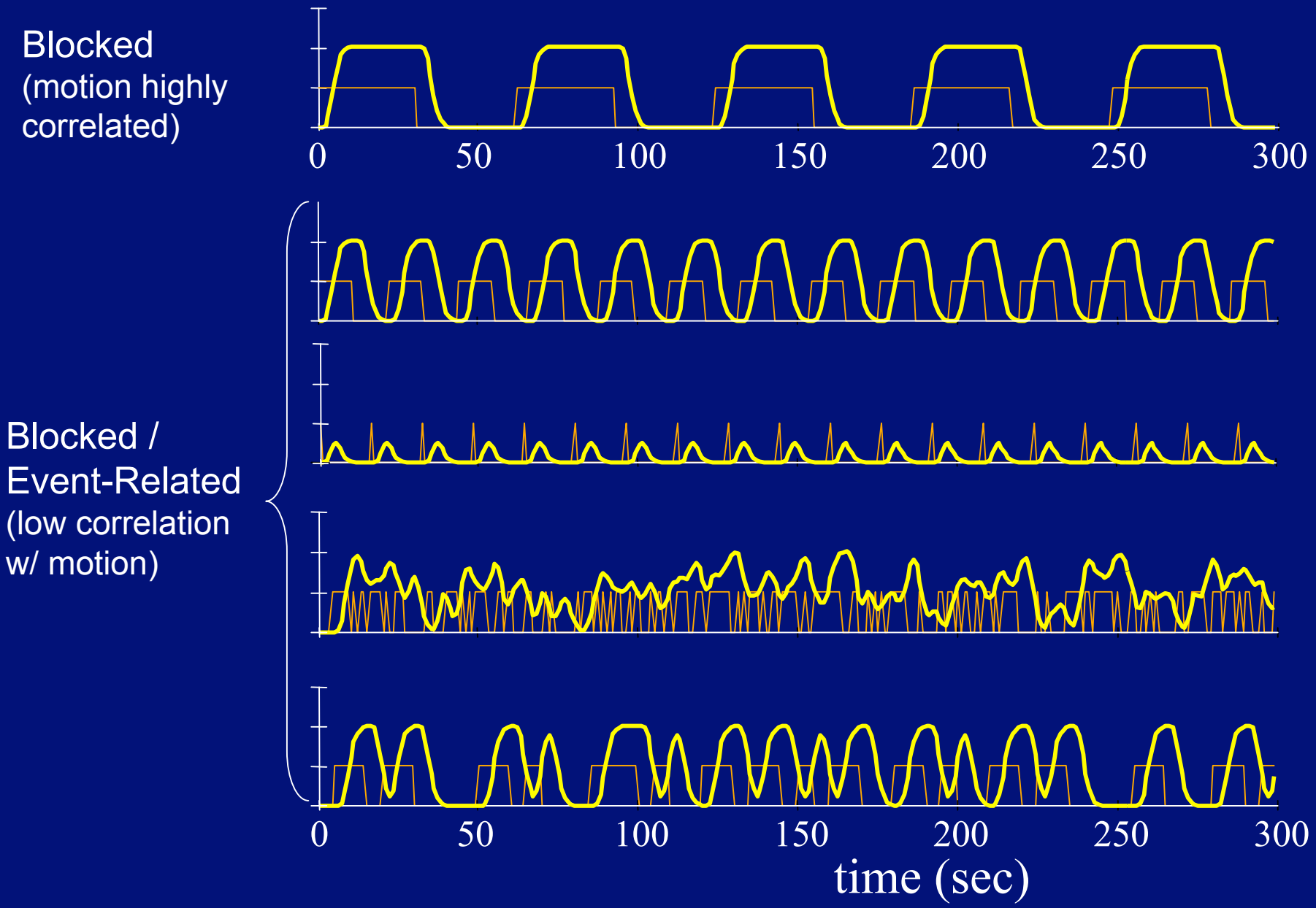


Event-Related Design



Slide from R. Birn

Avoid Motion by Optimizing Stimulus Timing



R.M. Birn, et al., NeuroImage, 23, 1046-1058, 2004.

Slide from R. Birn

Physiological Signal Monitoring

- Heart pulsation and breathing add variance to BOLD signals

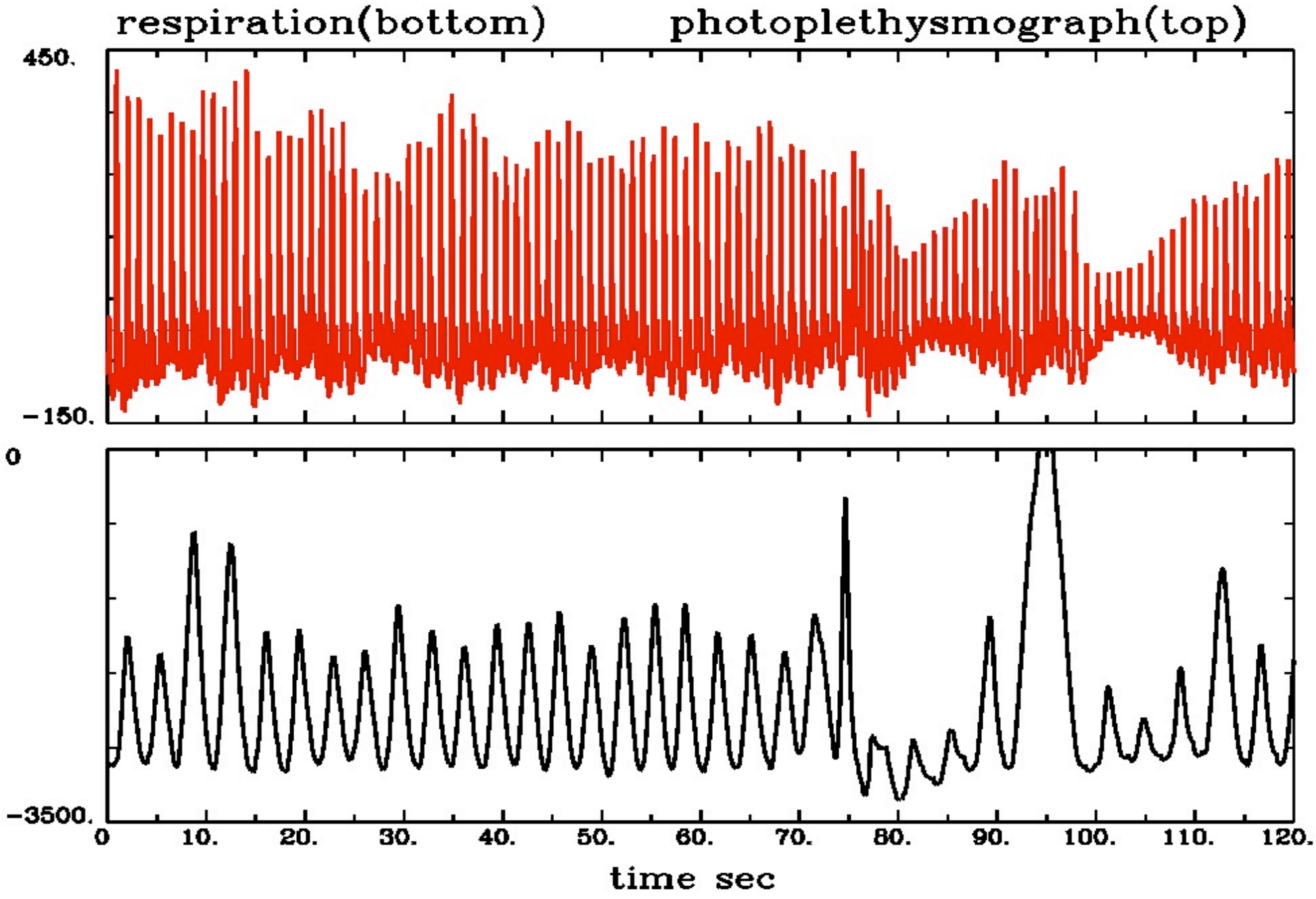
  Glover et al., Magn Reson Med 2000

  - Direct effects: movement of tissue and blood
  - Indirect effects: changes in baseline oxygenation

  Wise et al., NI 2004; Birn et al., NI 2006, 2008; Shmueli et al., NI 2007; Bianciardi et al., MRI 2009; Chang et al., NI 2009

- Modeling such effects reduces residual variance
  - At least 30% in majority of voxels
- These effects can be especially troublesome in resting state FMRI
  - Effects can be coherent over distant parts of cortex
  - Account for larger fraction of variance of interest

RETROICOR and RVT correction



Physiologic rate regressors: modeling issues

The phase and the shape of the expected fMRI signal changes due to fluctuations in the rates of respiratory and cardiac pulsation are not fully understood…

Respiration volume per unit time (RVT) regressor

$$\sim \frac{\Delta R}{\Delta t_R}$$

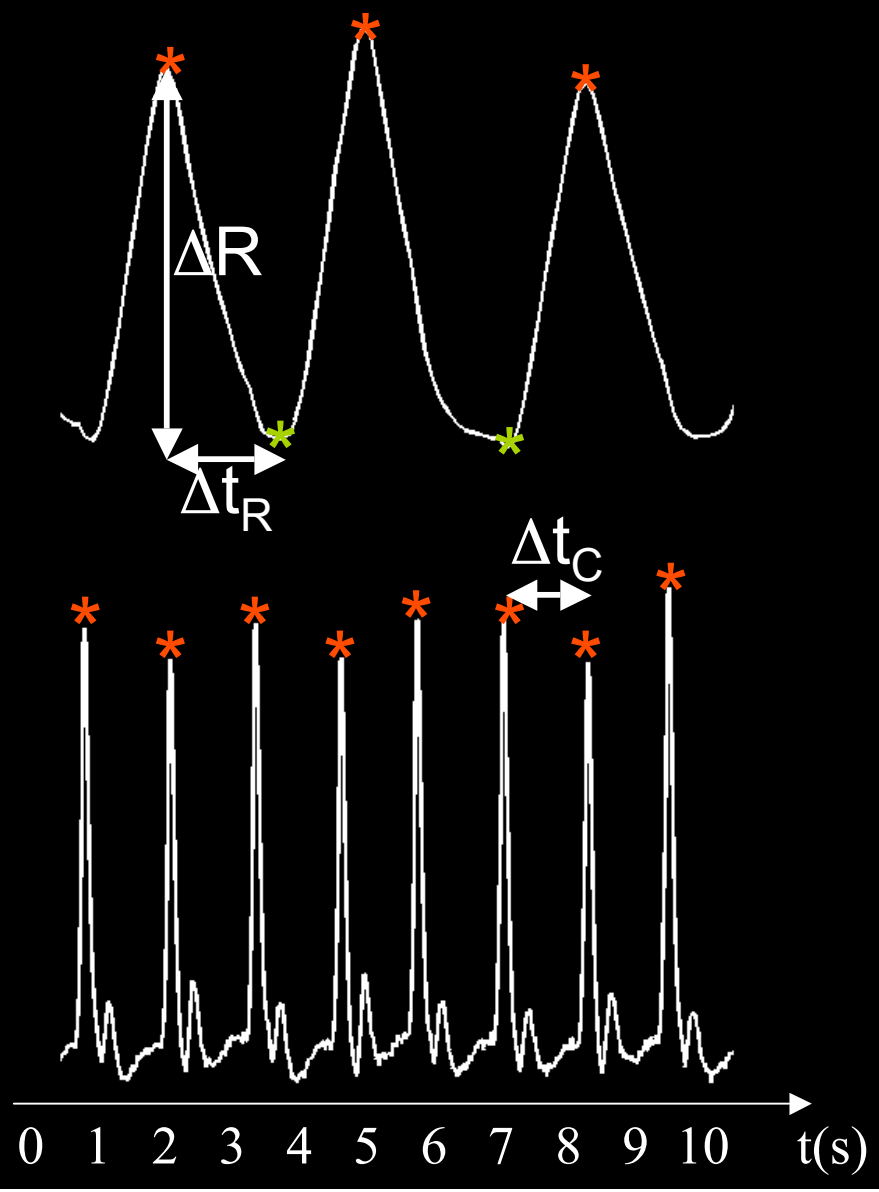
Birn et al., NI, 2006

Cardiac rate (CR) regressor

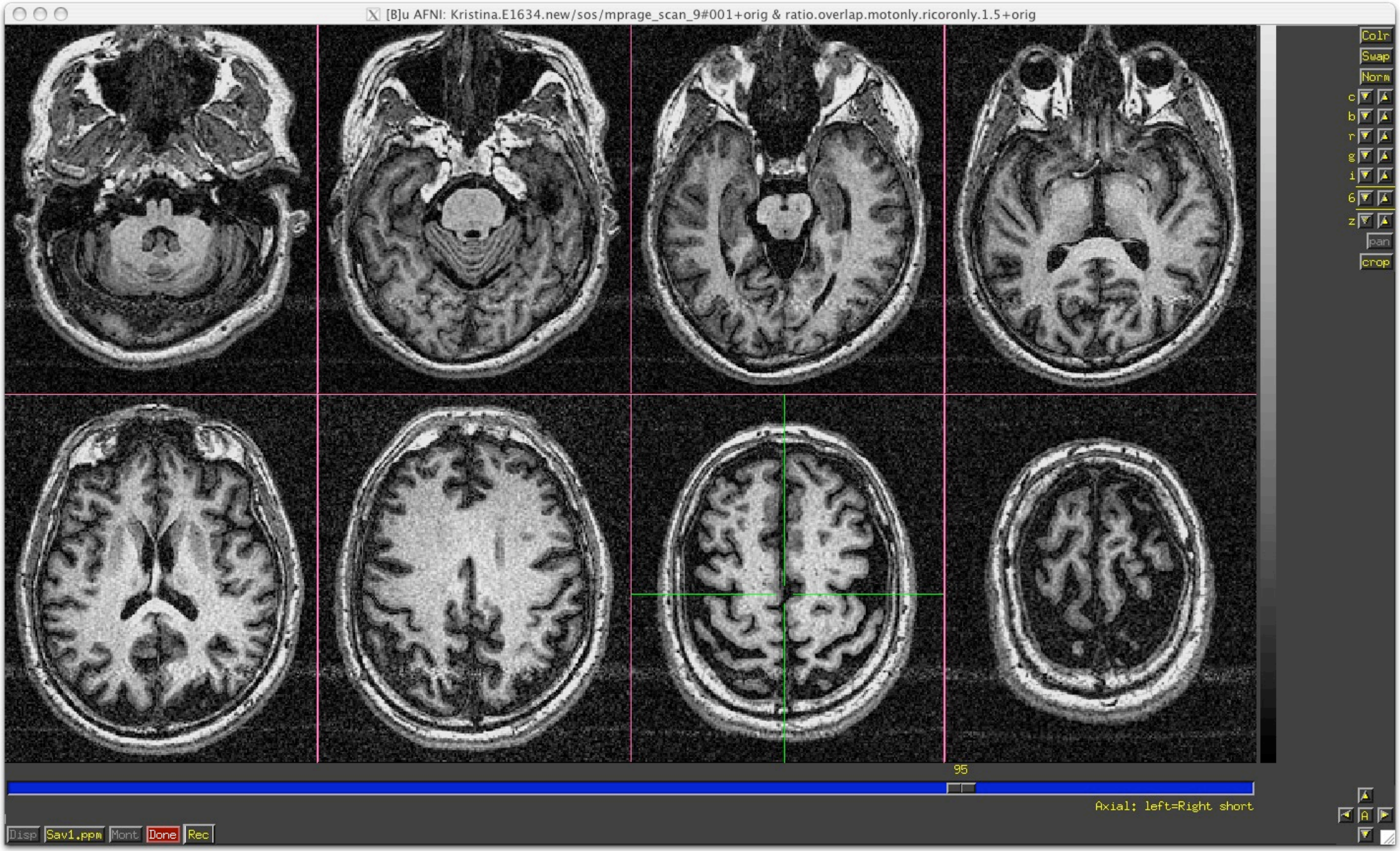
$$\sim \frac{1}{\Delta t_C}$$

Shmueli et al, NI, 2007

Slide courtesy of M. Bianciardi

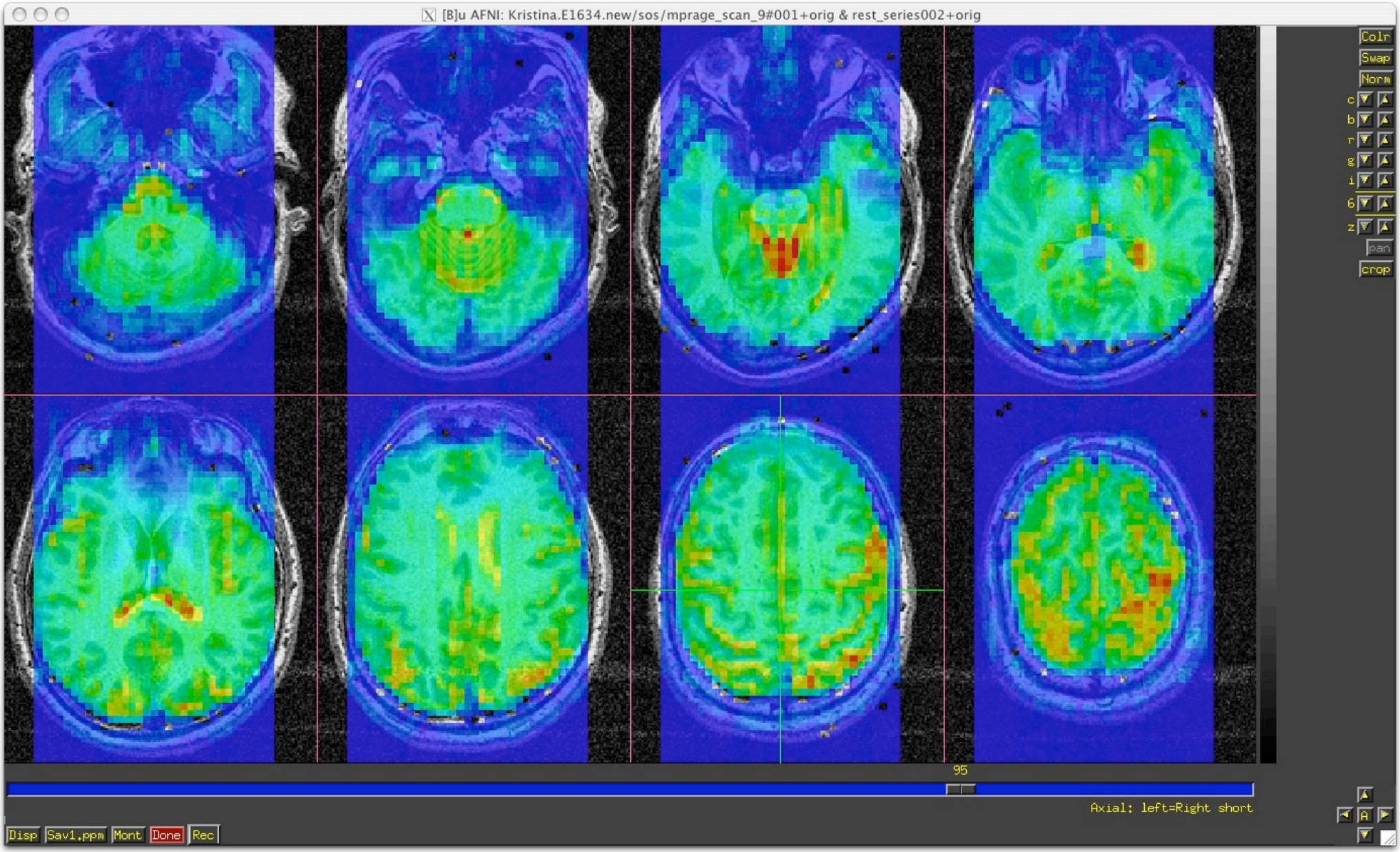


Sample R-RICOR Result: Anat



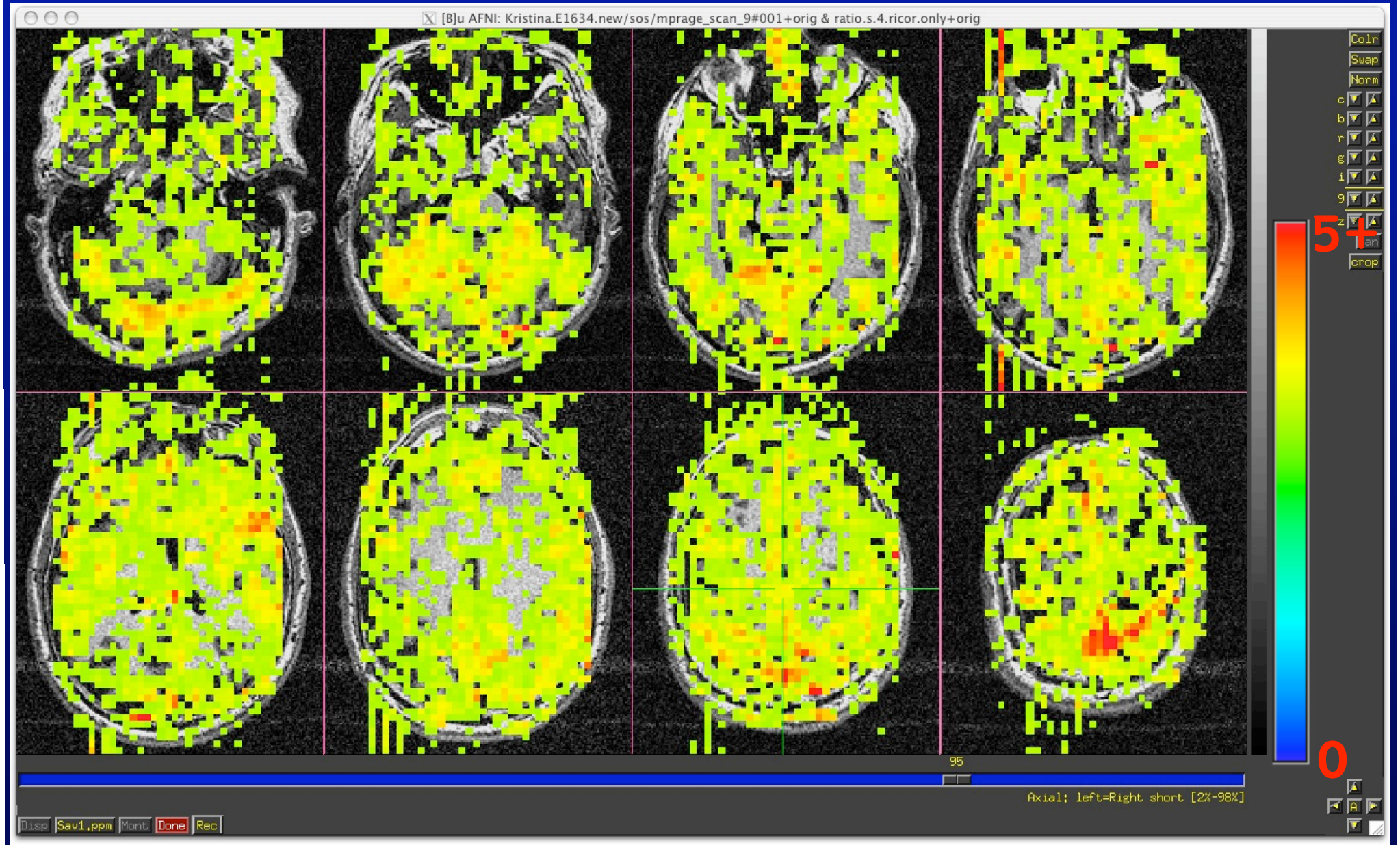
Data courtesy of K. Simonyan

Sample R-RICOR Result: Anat w/ EPI



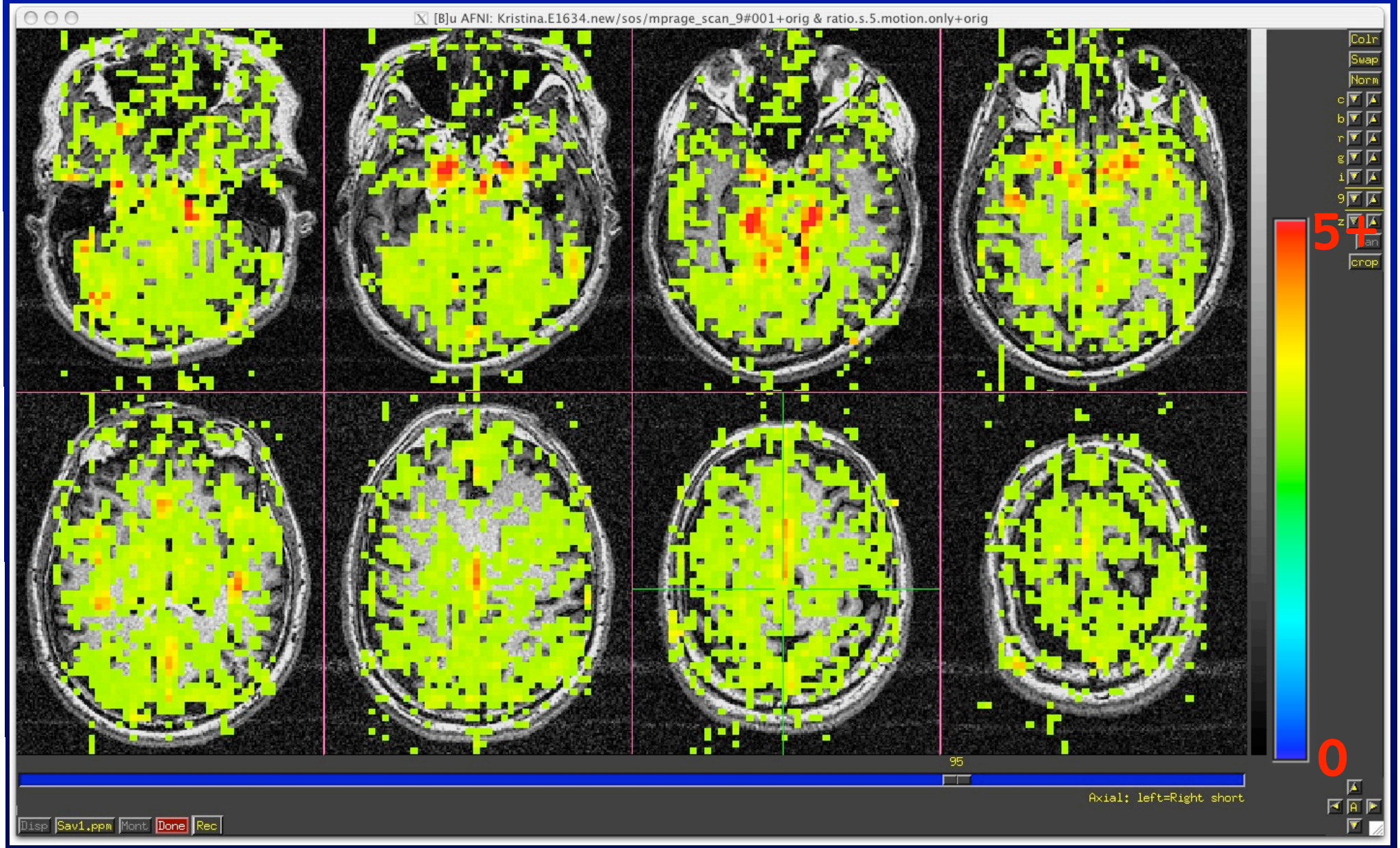
Data courtesy of K. Simonyan

SSE Ratio > 1.3: No Motion/Full Model



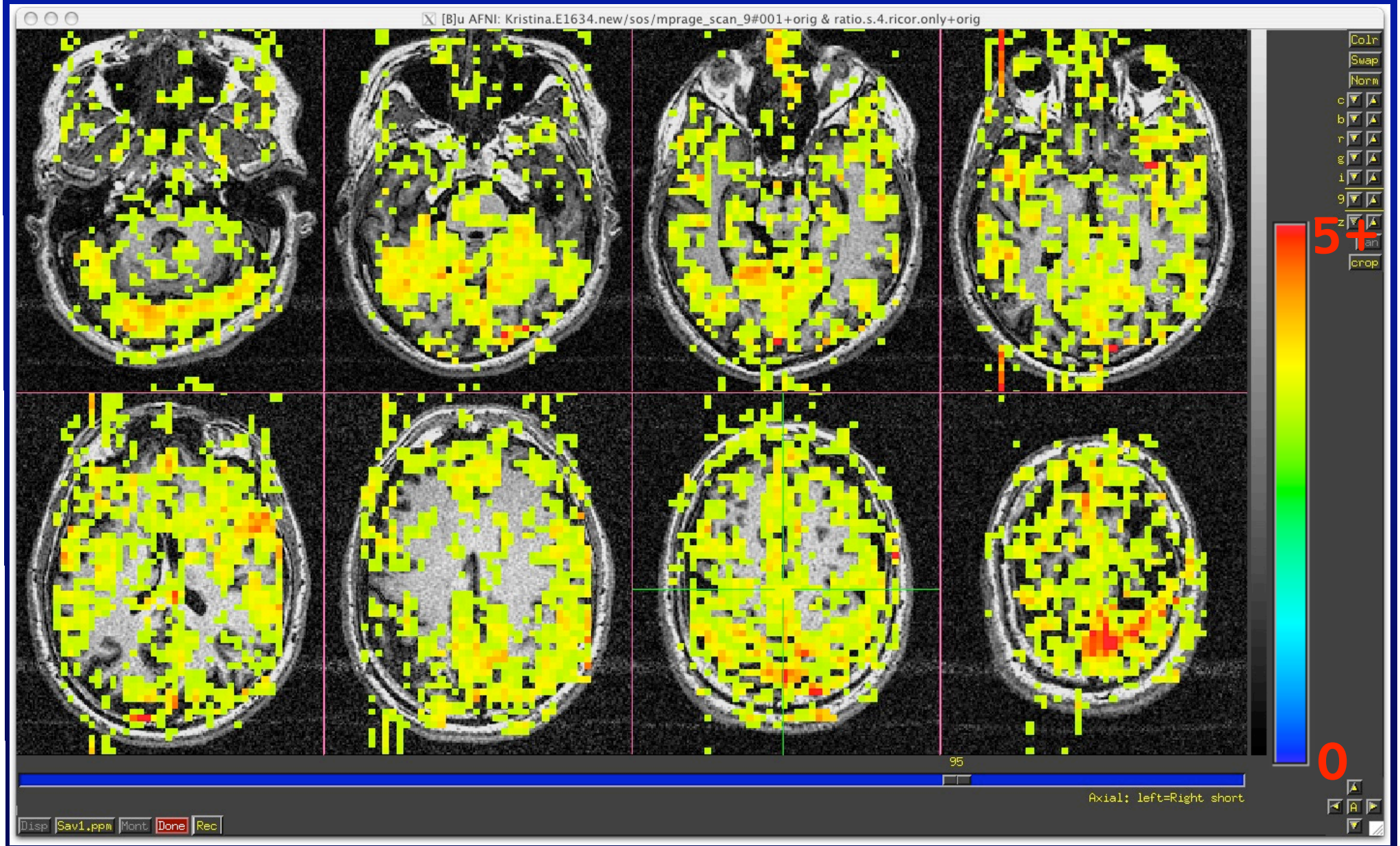
Data courtesy of K. Simonyan

SSE Ratio > 1.3: No R-RICOR/Full Model



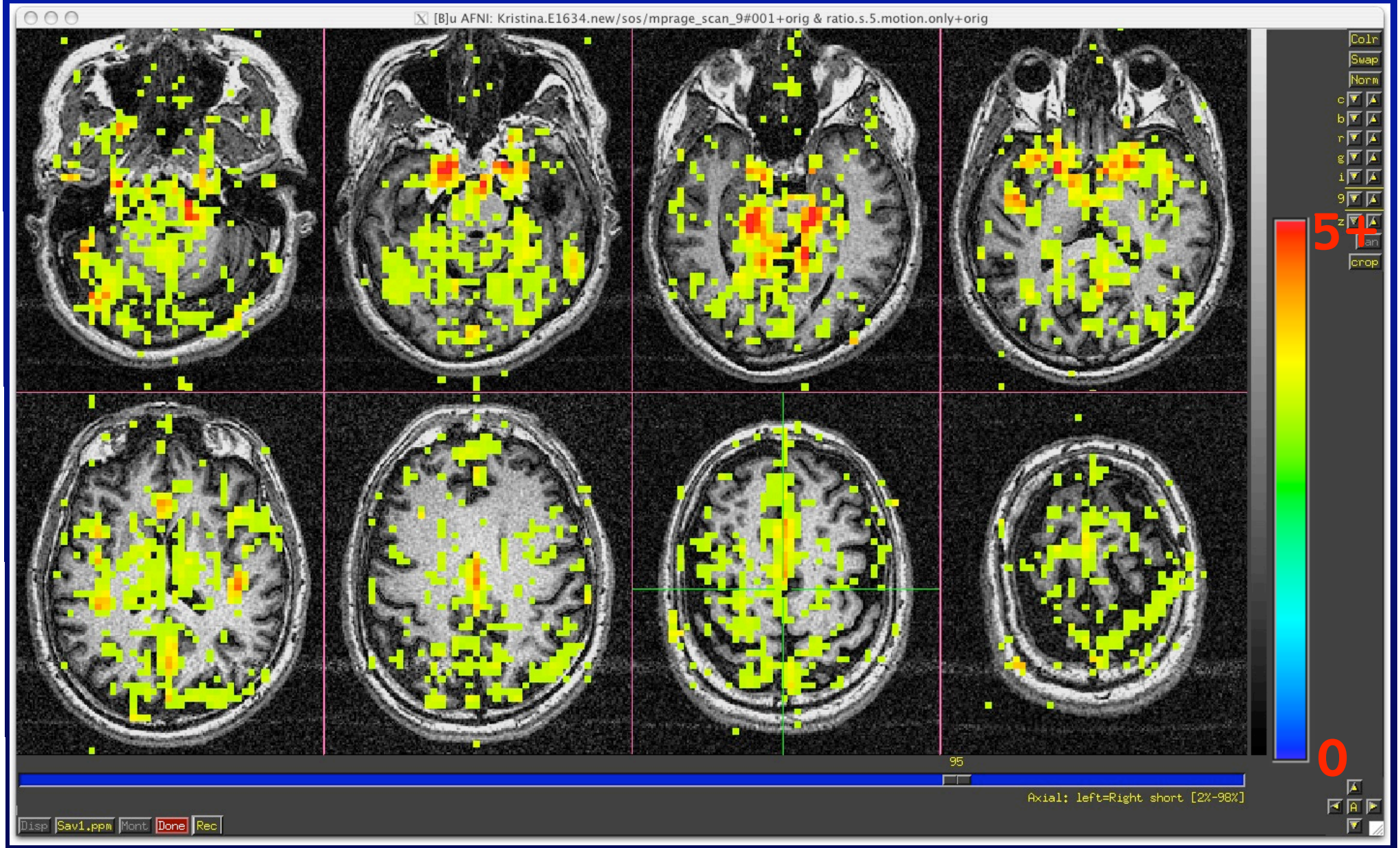
Data courtesy of K. Simonyan

SSE Ratio > 1.5: No Motion/Full Model



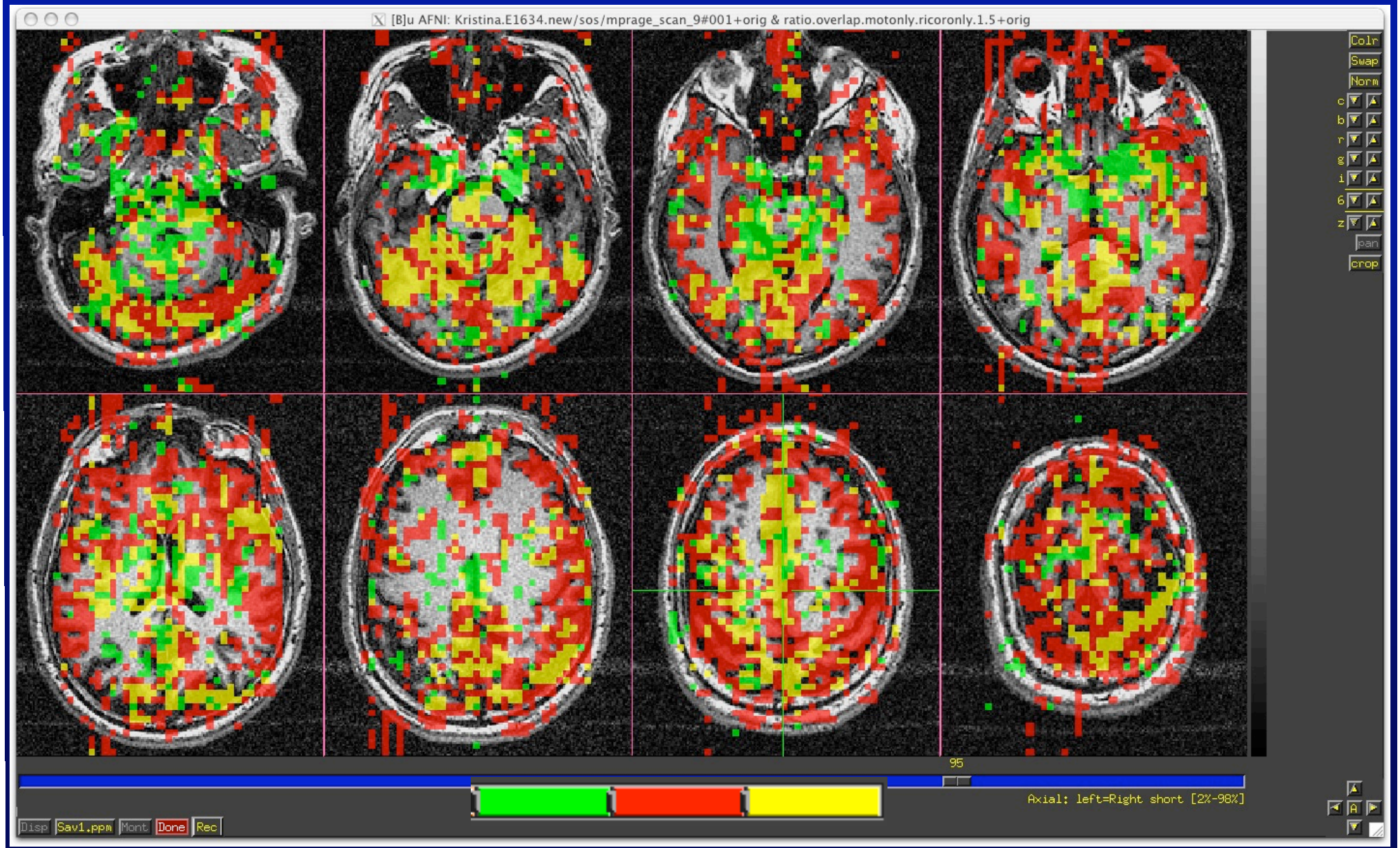
Data courtesy of K. Simonyan

SSE Ratio > 1.5: No R-RICOR/Full Model



Data courtesy of K. Simonyan

R-RICOR, Motion SSE Ratio Overlap



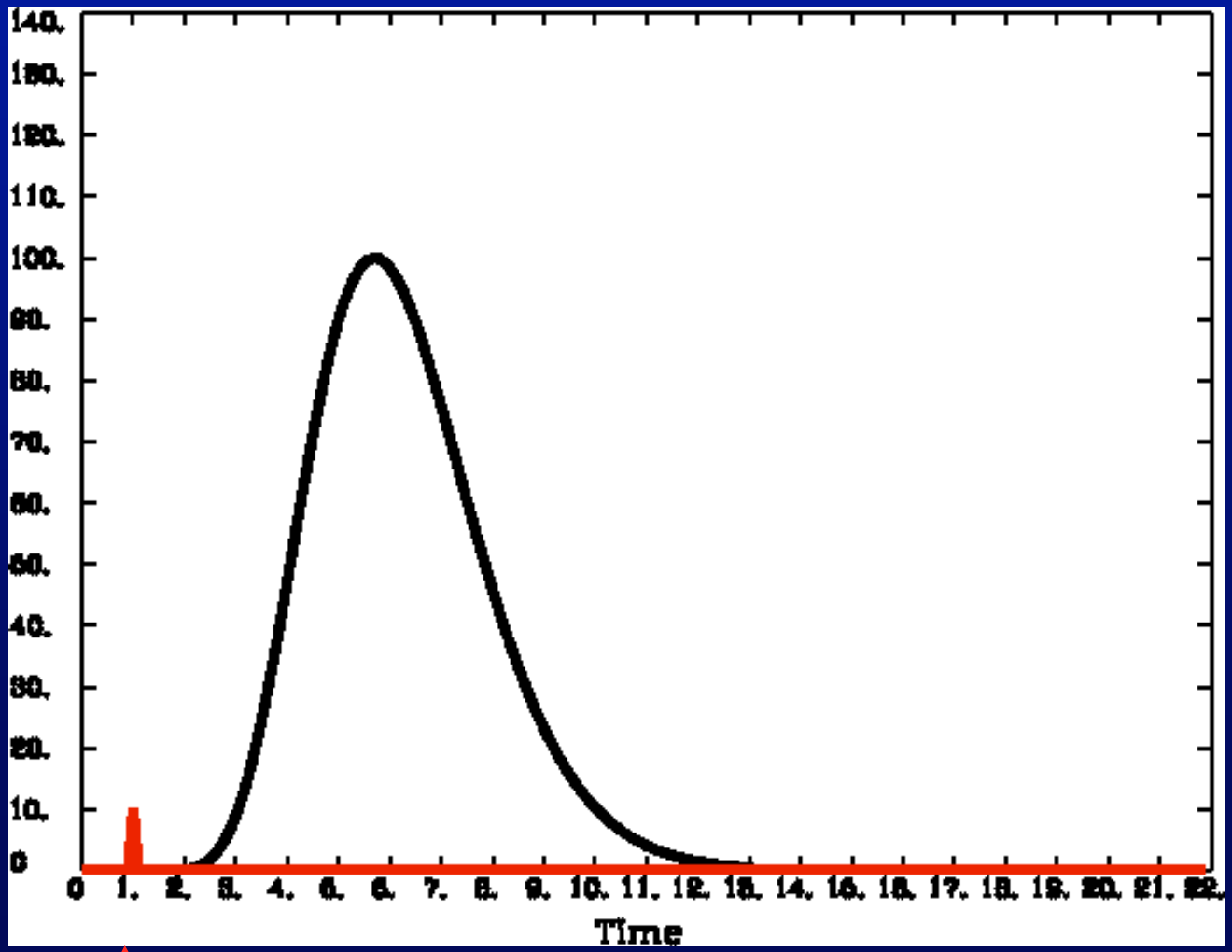
RR Mot both

Data courtesy of K. Simonyan

Linear Regression In Stage 1

Hemodynamic Response Function (HRF)

- **HRF** is the idealization of measurable FMRI signal change responding to a single activation cycle (up and down) from a stimulus in a voxel



Response to brief activation (< 1 s):

- delay of 1-2 s
- rise time of 4-5 s
- fall time of 4-6 s
- model equation:

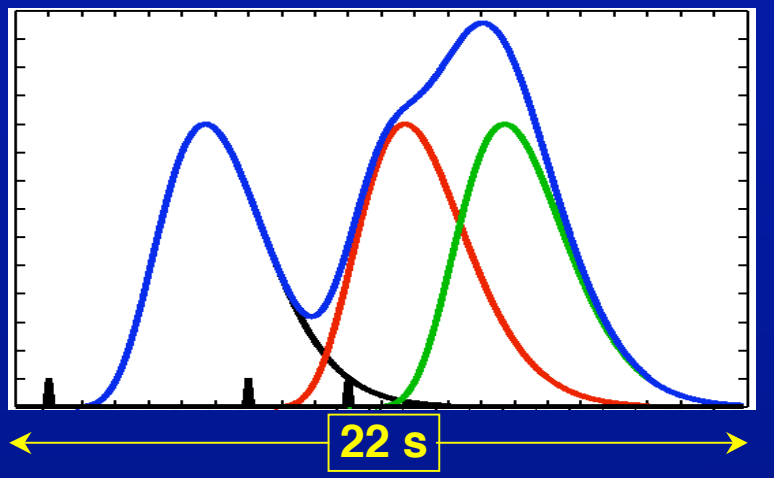
$$h(t) \propto t^b e^{-t/c}$$

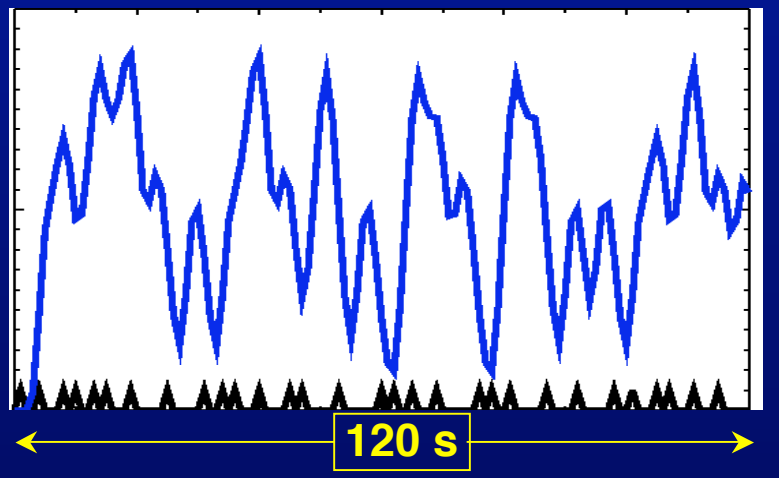
- $h(t)$ is signal change t seconds ***after*** activation

1 Brief Activation (Event)

Convolution Signal Model

- FMRI signal model (in each voxel) is taken as sum of the individual trial HRFs (assumed equal)
  - Stimulus timing is assumed known (or measured)
  - Resulting time series (in **blue**) are called the ***convolution*** of the HRF with stimulus timing
  - Finding HRF = "deconvolution"
  - AFNI code :`3dDeconvolve, 3dREMLfit`
  - Convolution models only the FMRI signal **changes**





- Real data starts at and returns to a nonzero, slowly drifting baseline

Simple Regression Models

- Assume a fixed shape $h(t)$ for the HRF
  - e.g., $h(t) = t^{8.6} \exp(-t/0.547)$ [MS Cohen, 1997]
  - Convolve with stimulus timing to get ideal response

$$r(t) = \sum_{k=1}^{K} h(t - \tau_k) = \text{sum of HRF copies}$$

- Assume a form for the baseline (data without activation)
  - e.g., $a + b \cdot t$ for a constant plus a linear trend
- In each voxel, fit data $Z(t)$ to a curve of the form

$$Z(t) \approx a + b \cdot t + \beta \cdot r(t)$$

The signal model*!*

  - a, b, β are unknown values to be found in each voxel
  - a, b are "nuisance" parameters
  - β is amplitude of $r(t)$ in data = "how much" BOLD
    - In this model, each stimulus assumed to get same BOLD response — in shape and in amplitude

Equations: Matrix-Vector Form

- Express ***known*** data vector as a sum of ***known*** columns with ***unknown*** coefficents:

$$\begin{bmatrix} z_0 \\ z_1 \\ z_2 \\ \vdots \\ z_{N-1} \end{bmatrix} \approx \begin{bmatrix} 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \cdot a + \begin{bmatrix} 0 \\ 1 \\ 2 \\ \vdots \\ N-1 \end{bmatrix} \cdot b + \begin{bmatrix} r_0^{(1)} \\ r_1^{(1)} \\ r_2^{(1)} \\ \vdots \\ r_{N-1}^{(1)} \end{bmatrix} \cdot \beta_1 + \begin{bmatrix} r_0^{(2)} \\ r_1^{(2)} \\ r_2^{(2)} \\ \vdots \\ r_{N-1}^{(2)} \end{bmatrix} \cdot \beta_2 + \cdots$$

- Const baseline
- Linear trend
- Response to stim#1
- Response to stim#2

'≈' is "least squares"

or

$$\begin{bmatrix} z_0 \\ z_1 \\ z_2 \\ \vdots \\ z_{N-1} \end{bmatrix} \approx \begin{bmatrix} 1 & 0 & r_0^{(1)} & r_0^{(2)} & \cdots \\ 1 & 1 & r_1^{(1)} & r_1^{(2)} & \cdots \\ 1 & 2 & r_2^{(1)} & r_2^{(2)} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \\ 1 & N-1 & r_{N-1}^{(1)} & r_{N-1}^{(2)} & \cdots \end{bmatrix} \begin{bmatrix} a \\ b \\ \beta_1 \\ \beta_2 \\ \vdots \end{bmatrix}$$

or

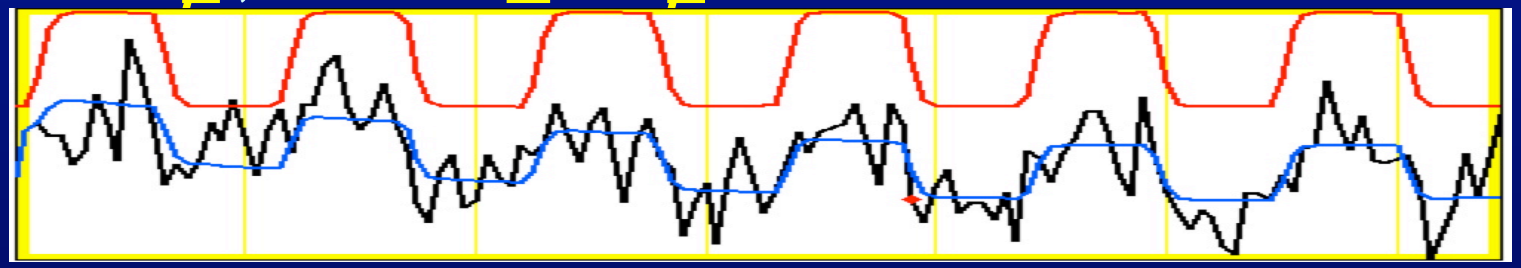
$$\underbrace{\mathbf{z}}_{\text{vector of data}} \approx \underbrace{\mathbf{R}}_{\text{matrix of columns}} \underbrace{\beta}_{\text{vector of coeff}}$$

the "design" matrix; AKA **X**

z depends on the voxel; **R** doesn't

Solving $z \approx R\beta$ for β

- Number of equations = number of time points
  - 100s per run, but perhaps 1000s per subject
- Number of unknowns usually in range 5–50
- Least squares solution: $\underline{\beta} = [R^T R]^{-1} R^T z$
  - $\underline{\beta}$ denotes an *estimate* of the true (unknown) β
  - From $\underline{\beta}$, calculate $\underline{z} = R\,\underline{\beta}$ as the *fitted model*



- $z - \underline{z}$ is the ***residual time series*** = noise (we hope)
- Statistics measure how much each regressor helps reduce residuals

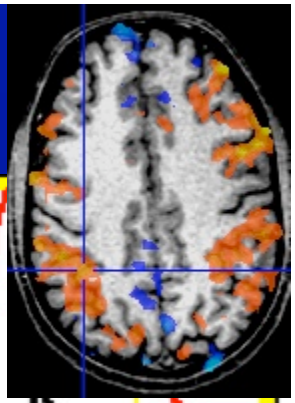
- Collinearity: when matrix $R^T R$ can't be inverted
  - Near collinearity: when inverse exists but is huge

Simple Regression: Recapitulation

- Choose HRF model $h(t)$ [AKA *fixed-model regression*]
- Build model responses $r_n(t)$ to each stimulus class
  - Using $h(t)$ and the stimulus timing
- Choose baseline model time series
  - Constant + linear + quadratic (+ movement?)
- Assemble model and baseline time series into the columns of the $\mathbf{R}$ matrix
- For each voxel time series $\mathbf{z}$, solve $\mathbf{z} \approx \mathbf{R}\beta$ for $\underline{\beta}$
- **Individual subject maps**: Test the coefficients in $\underline{\beta}$ that you care about for statistical significance
- **Group maps**: Transform the coefficients in $\underline{\beta}$ that you care about to Talairach space, and perform statistics on the collection of $\underline{\beta}$ values across subjects

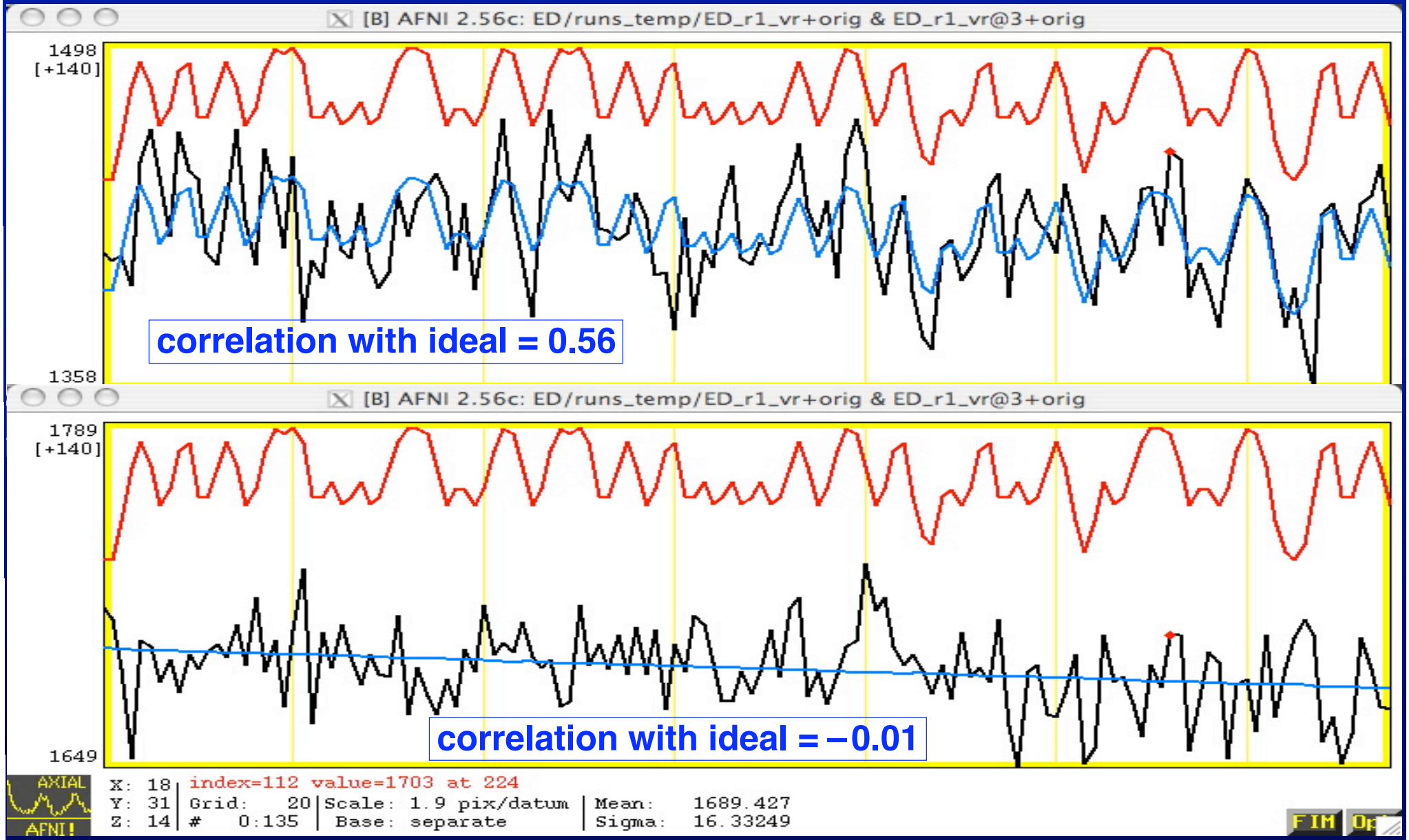
Take a look at your model fit

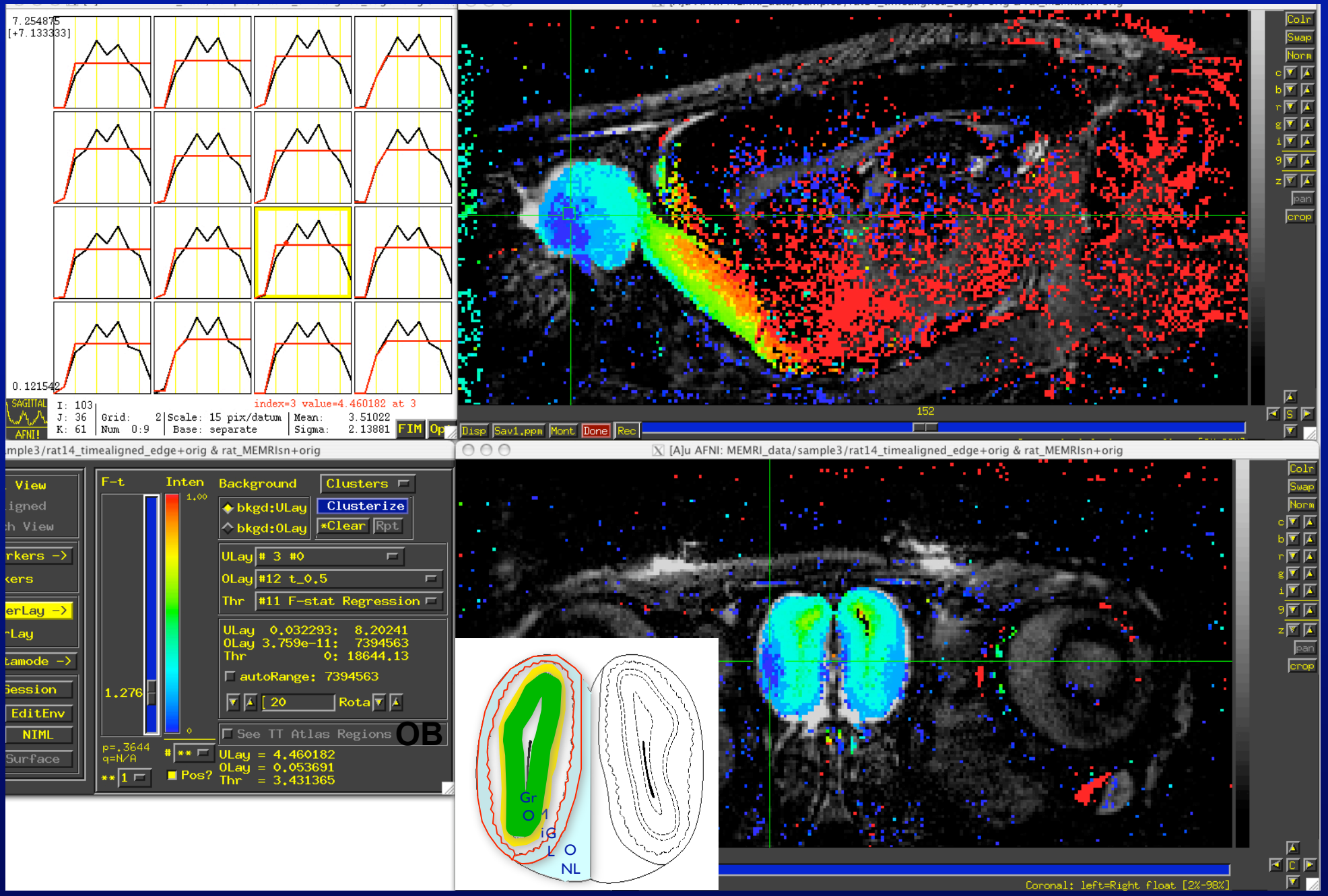
Same Voxel: Runs 1 and 2



Block-trials: 27 s “on” / 27 s “off”; TR=2.5 s; 130 time points/run

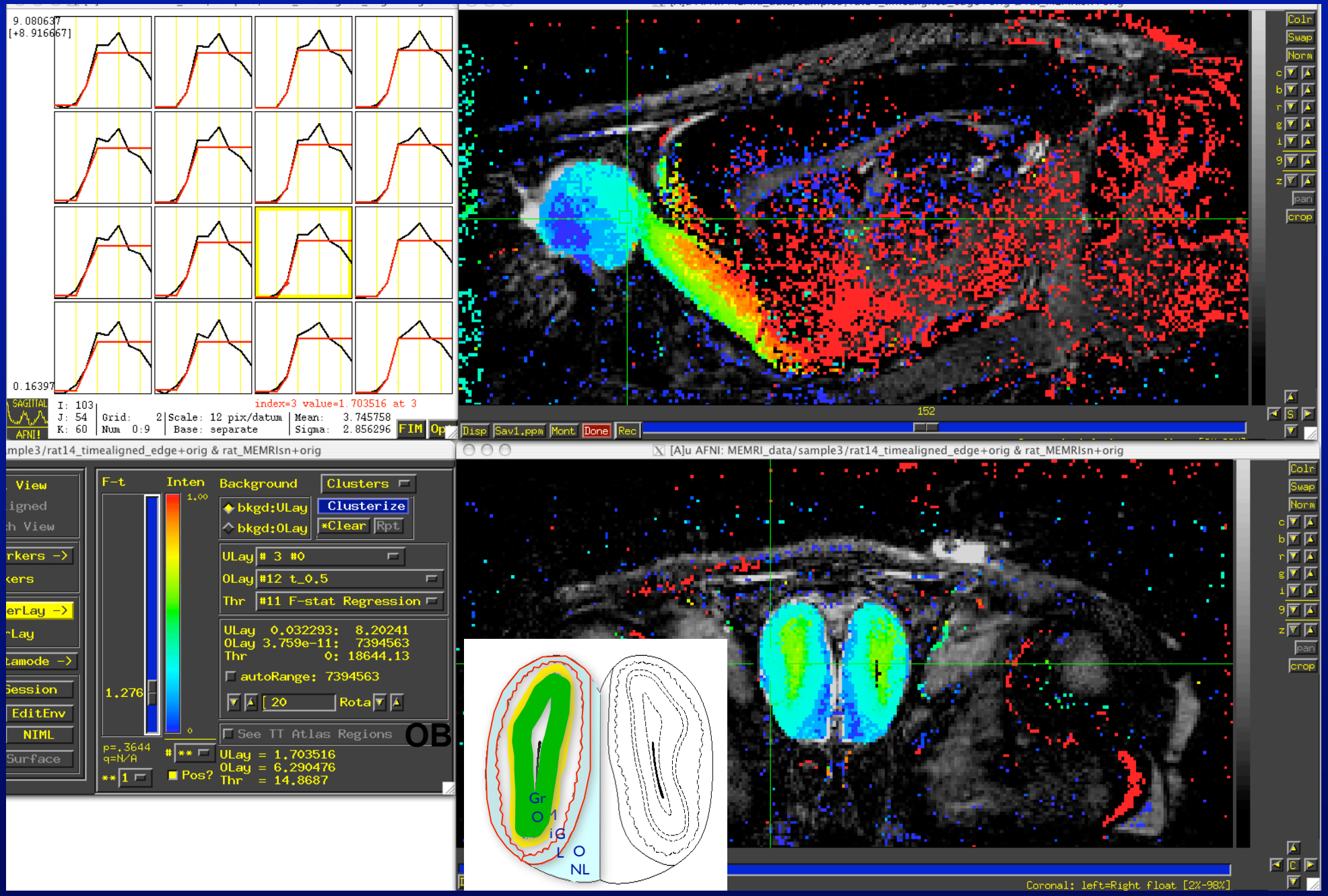
Two Voxel Time Series from Same Run





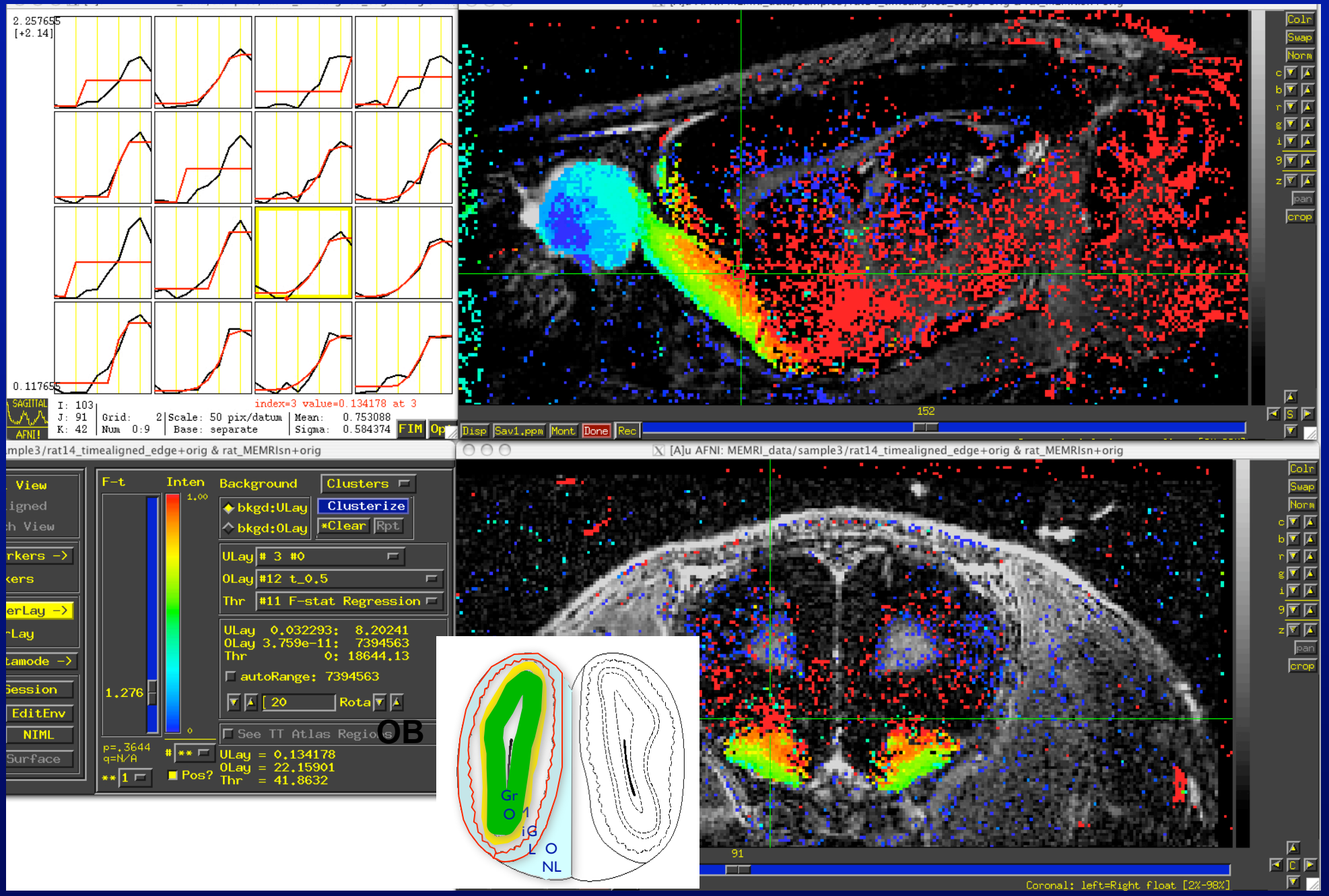
Data courtesy of Der-Yow Chen & Alan Koretsky, NINDS/NIH

Z.S.S 8-09



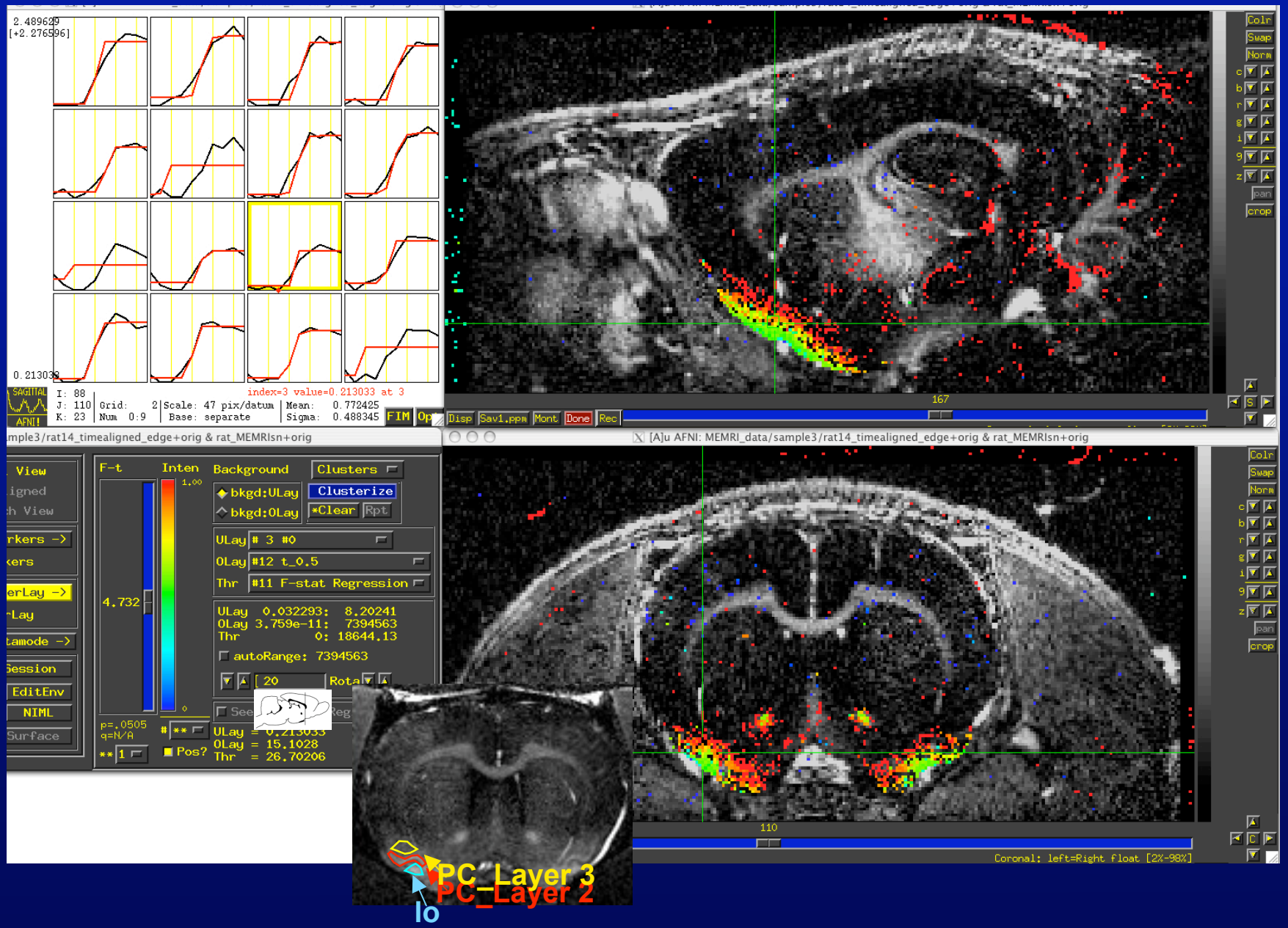
Data courtesy of Der-Yow Chen & Alan Koretsky, NINDS/NIH

Z.S.S 8-09



Data courtesy of Der-Yow Chen & Alan Koretsky, NINDS/NIH

Z.S.S 8-09



Data courtesy of Der-Yow Chen & Alan Koretsky, NINDS/NIH

Deconvolution Signal Models

- Simple or Fixed-shape regression :
  - We fixed the shape of the HRF — amplitude varies
  - Used **`-stim_times`** to generate the signal model (AKA the "ideal") from the stimulus timing
  - Found the amplitude of the signal model in each voxel — solution to the set of linear equations = β weights
- Deconvolution or Variable-shape regression :
  - We allow the shape of the HRF to vary in each voxel, for each stimulus class
  - Appropriate when you don't want to over-constrain the solution by assuming an HRF shape
  - ***Caveat***: need to have enough time points during the HRF in order to resolve its shape (e.g., TR ≤ 3 s)

Deconvolution: Pros & Cons (+ & –)

+ Letting HRF shape varies allows for subject and regional variability in hemodynamics
+ Can test HRF estimate for different shapes (e.g., are later time points more "active" than earlier?)
+ Weird shapes in HRF usually indicate problem with timing, design, etc.
– Need to estimate more parameters for each stimulus class than a fixed-shape model (e.g., 4-15 *vs*. 1 parameter = amplitude of HRF)
– Which means you need more data to get the same statistical power (assuming that the fixed-shape model you would otherwise use was in fact "correct")

Expressing HRF via Regression Unknowns

- The tool for expressing an unknown function as a finite set of numbers that can be fit via linear regression is an **expansion in basis functions**

$$h(t) = \beta_0\psi_0(t) + \beta_1\psi_1(t) + \beta_2\psi_2(t) + \cdots = \sum_{q=0}^{q=p} \beta_q\psi_q(t)$$

  - The basis functions $\psi_q(t)$ & expansion order p are known
    - Larger $p \Rightarrow$ more complex shapes & more parameters
  - The unknowns to be found (in each voxel) comprises the set of weights β_q for each $\psi_q(t)$

- β weights appear only by multiplying known values, and HRF only appears in signal model by linear convolution (addition) with known stimulus timing
  - Resulting signal model still solvable by linear regression

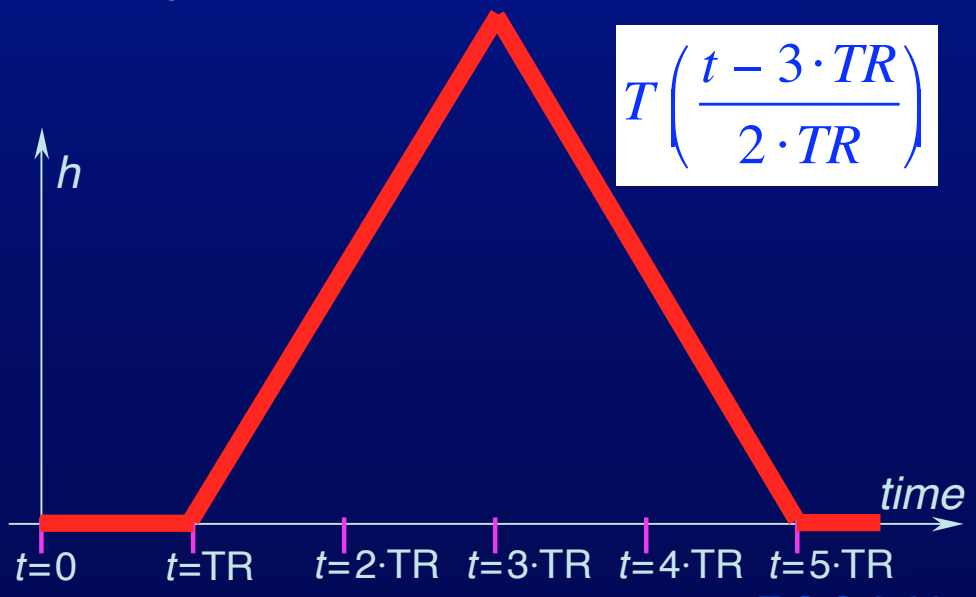
3dDeconvolve with "Tent Functions"

- Need to describe HRF shape and magnitude with a finite number of parameters
  - And allow for calculation of $h(t)$ at any arbitrary point in time after the stimulus times:

$$r_n = \sum_{k=1}^{K} h(t_n - \tau_k) = \text{sum of HRF copies}$$

- Simplest set of such functions are **tent functions**
  - Also known as "piecewise linear splines"

$$T(x) = \begin{cases} 1 - |x| & \text{for } -1 < x < 1 \\ 0 & \text{for } |x| > 1 \end{cases}$$

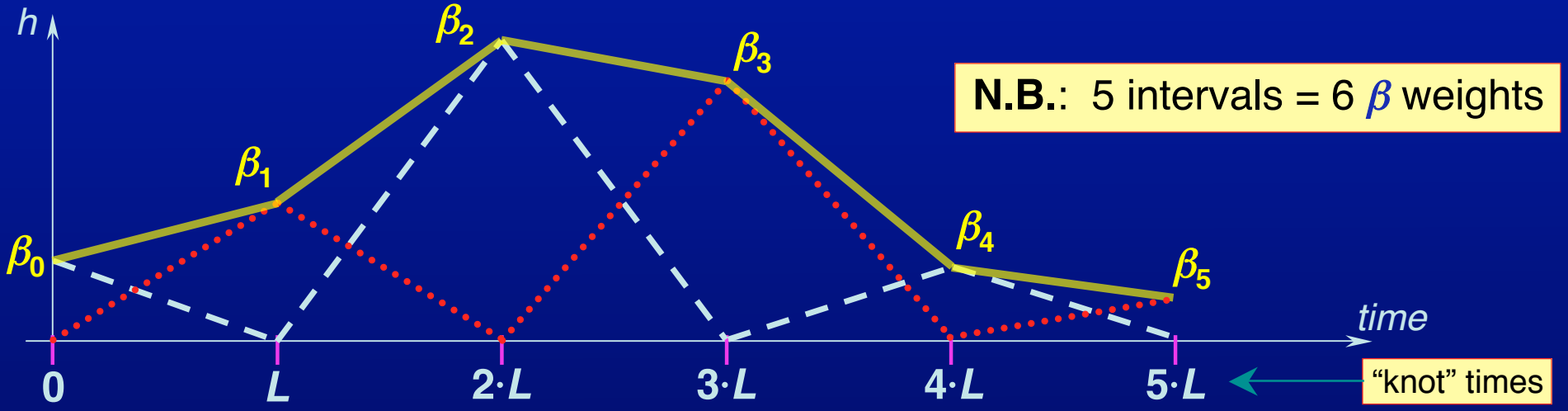


Tent Functions = Linear Interpolation

A

- Expansion of HRF in a set of spaced-apart tent functions is the same as linear interpolation between "knots"

$$h(t) = \beta_0 \cdot T\left(\frac{t}{L}\right) + \beta_1 \cdot T\left(\frac{t-L}{L}\right) + \beta_2 \cdot T\left(\frac{t-2\cdot L}{L}\right) + \beta_3 \cdot T\left(\frac{t-3\cdot L}{L}\right) + \cdots$$



N.B.: 5 intervals = 6 β weights

- Tent function parameters are also easily interpreted as function values (e.g., β_2 = response at time $t = 2 \cdot L$ after stim)
- User must decide on relationship of tent function grid spacing L and time grid spacing TR (usually would choose $L \geq$ TR)
- In `3dDeconvolve/3dREMLfit`: specify duration of HRF and number of β parameters

Deconvolution Regression

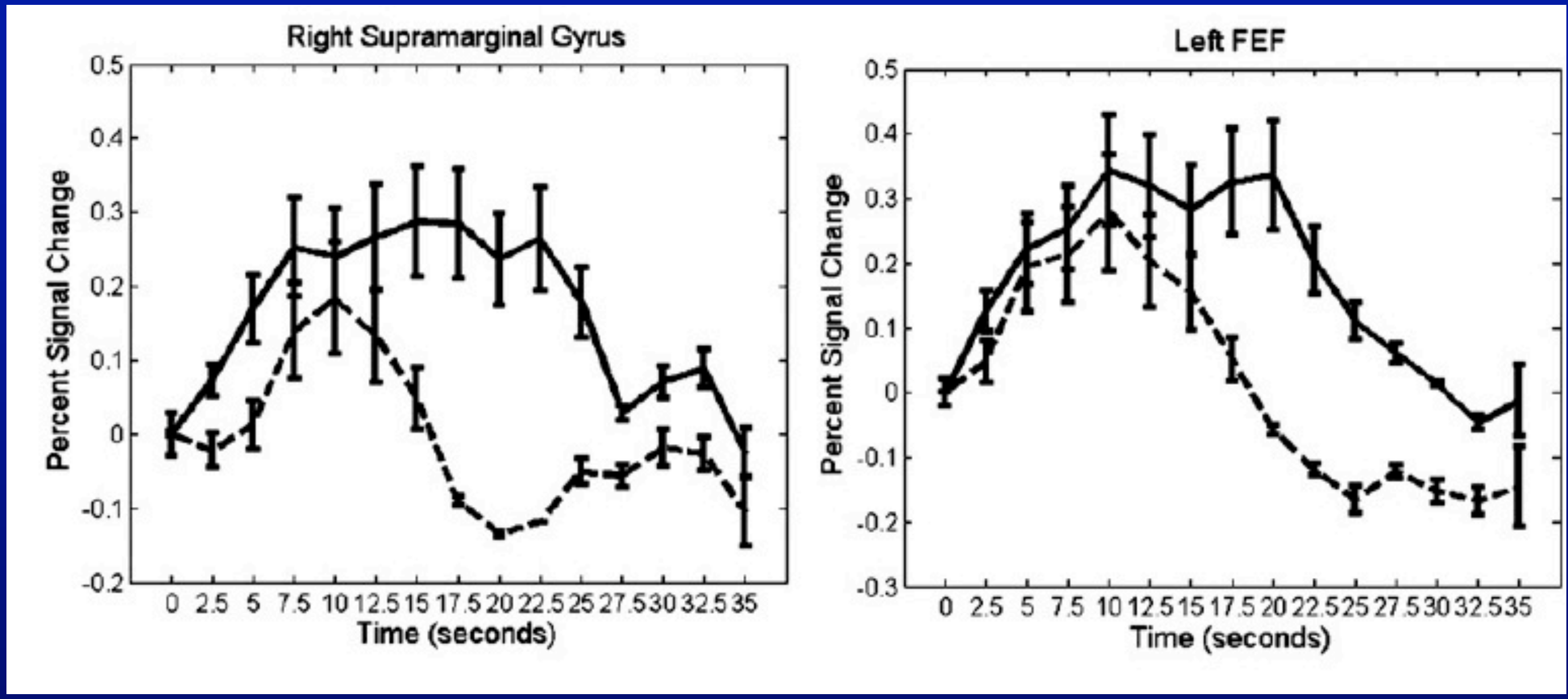


Fig. 4 From Geier C.F. et al. NI07

AM Regression - 1

- **AM** = **A**mplitude **M**odulated (or Modulation)
  - Have some extra data measured about each response to a stimulus, and *maybe* the BOLD response amplitude is modulated by this
  - Reaction time; Galvanic skin response; Pain level perception; Emotional valence (happy or sad or angry face?)
- Want to see if some brain activations is linearly proportionally to one or more **ABI** (**A**uxiliary **B**ehaviorial **I**nformation)

- Need to make 2 separate regressors
  - One to find the mean FMRI response (the usual `-stim_times` analysis)
  - One to find the variations in the FMRI response as the ABI data varies
- The second regressor should have the form

$$r_{\mathrm{AM2}}(t) = \sum_{k=1}^{K} h(t-\tau_k)\cdot(a_k - \bar{a})$$

  - Where a_k = value of k^{th} ABI value, and a is the average ABI value
- Statistics and β for second regressor make activation map of places whose BOLD response changes with changes in ABI

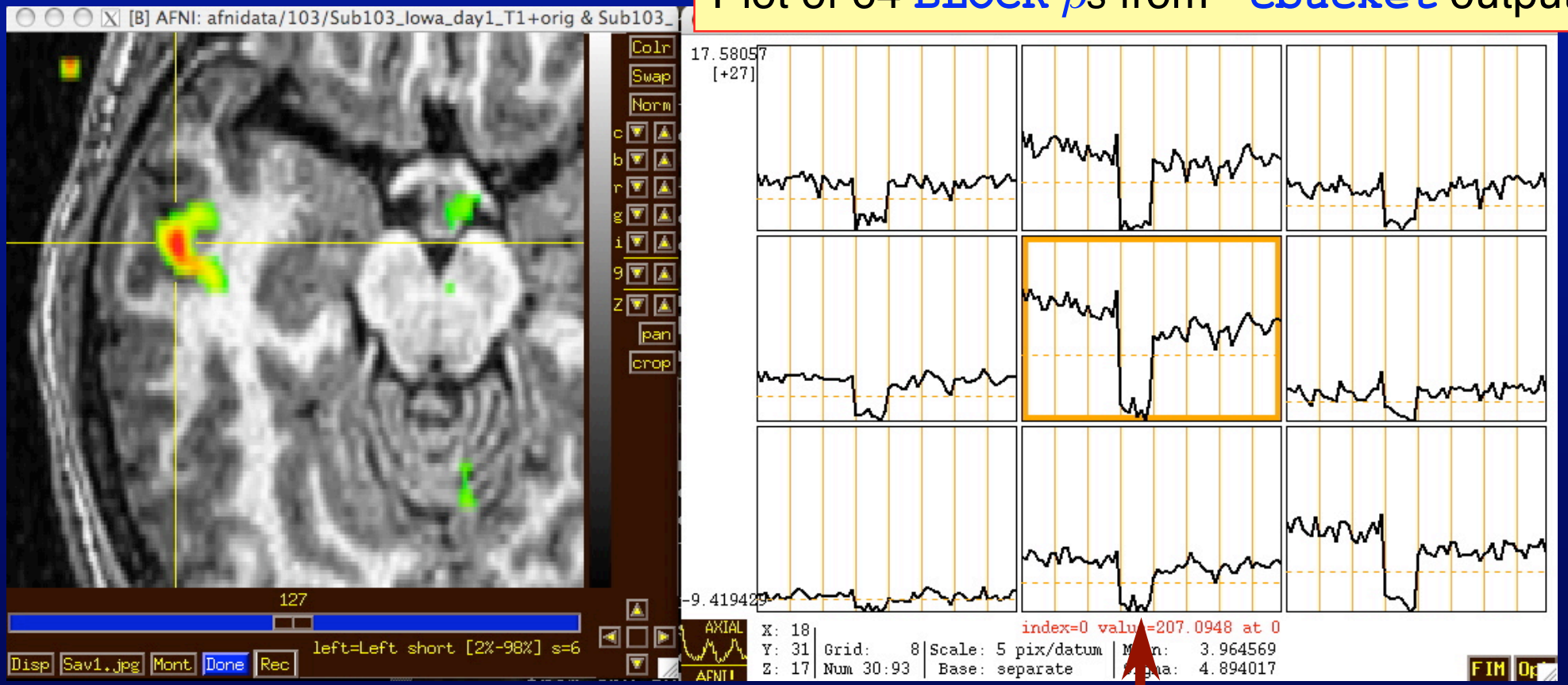
IM Regression - 1

- **IM** = **I**ndividual **M**odulation
  - –Compute *separate* amplitude of response for each stimulus
    - •Instead of computing average amplitude of responses to multiple stimuli in the same class
  - –Response amplitudes (βs) for each individual block/event will be highly noisy
    - •Can't use individual activation map for much
    - •Must pool the computed βs in some further statistical analysis (*t*-test? inter-voxel correlations in the βs? model βs as a function of some stimulus parameter ?)
  - –Usage: **`-stim_times_IM k tname model`**
    - •Like **`-stim_times`**, but creates a separate regression matrix column for each time given

IM Regression - 2

- IM estimates over 64 stimulus events
- Experiment: 64 blocks of sensorimotor task (8 runs each with 8 blocks)
- No exciting trend there, but notice sign reversal due to protocol error.

Plot of 64 **BLOCK** βs from **-cbucket** output



N.B.: sign reversal in run #4 = stimulus timing error!

Variance and serially correlated noise

- White noise estimate of variance:
  - N = number of time points; i = time index
  - m = number of fit parameters
  - $N-m$ = degrees of freedom (DOF) = how many equal-variance independent random values are left after the time series is fit with m regressors
    - •OLSQ assumption is that each of the N noise values in the data time series are equal-variance and independent (AKA white noise)

$$\hat{\sigma}^2 = \frac{1}{N-m}\sum_{i=0}^{N-1}[\text{data}_i - \text{fit}_i]^2$$

- If noise values *aren't* independent, then $N-m$ is too large an estimate of DOF, so variance estimate is too small
- Two possible solutions are:

1)Adjust variance estimate (and so the t- and F-values) to allow for fewer DOF

2)Come up with a different variance estimator that has all $N-m$ DOF possible (`3dREMLfit`)

- o Requires estimating the temporal correlation structure of the noise as well
- o Once temporal correlation matrix is known, use Generalized Least Squares (GLSQ; AKA pre-whitening) to estimate β parameter vector
- o GLSQ is consistent and should produce β with smaller variance than OLSQ

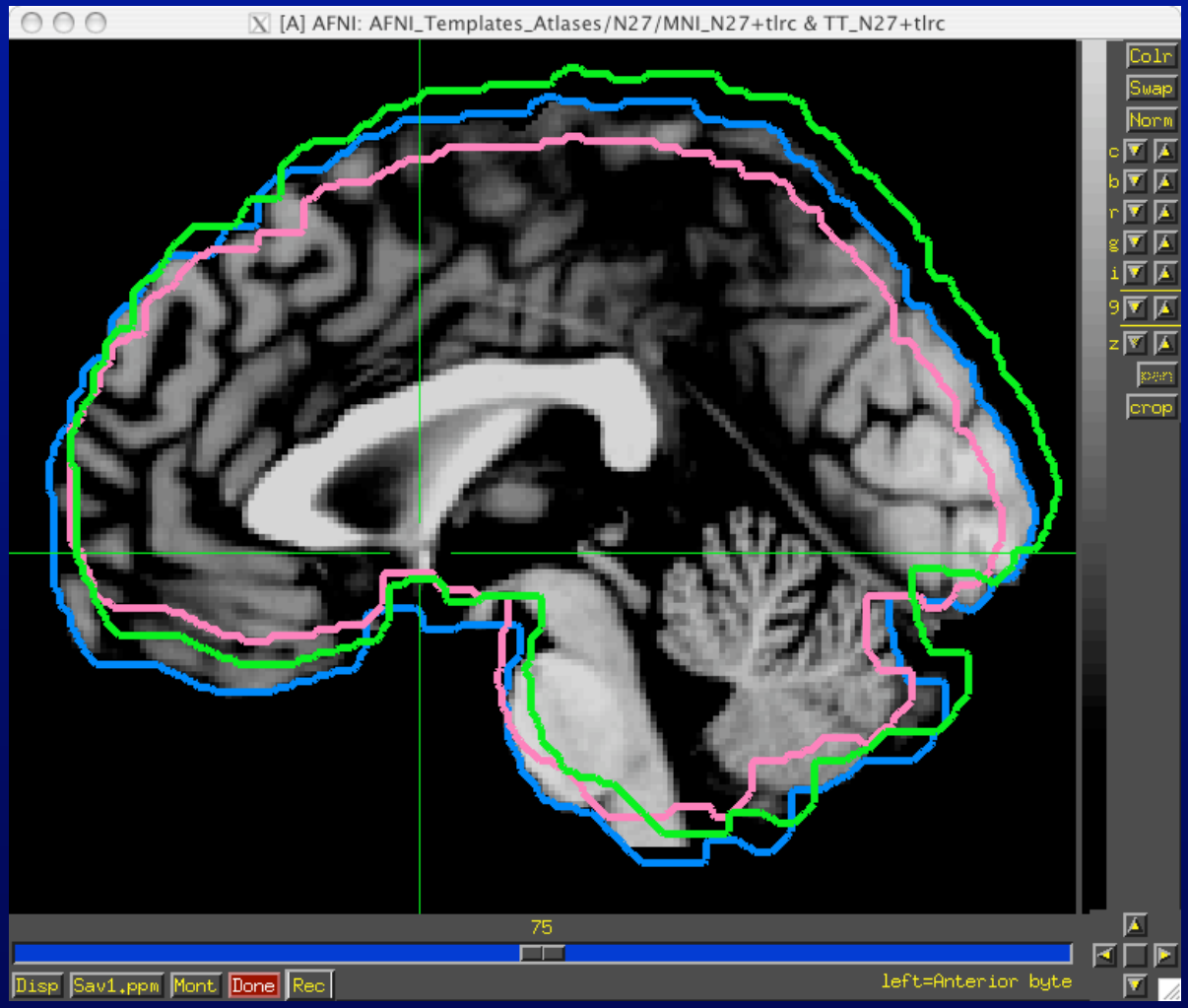
OLSQ and GLSQ

- Activation magnitudes (β s) estimated using OLSQ (Ordinary Least Squares) are consistent
  - "Consistent" means that if you repeated the identical experiment infinitely many times, and averaged the estimated value (*e.g.*, β ; σ^2), result would be the true value
- Variance of β s (σ^2s) is under-estimated with OLSQ in the presence of serial correlation
  - If the variance is under-estimated, then the individual subject t- and F-statistics will be over-estimated
- Group (stage 2) models that ignore (σ^2s) (ttest, anova, etc) give same results as models whether β s we obtained from OLSQ or GLSQ.
  - However newer approaches do carry β s *and* σ^2s to group statistics so GLSQ is needed

Keith Worsley's FMRIstat, FSL's FLAME AFNI's 3dMEMA

Standard Space For Group Analysis

- Single subject data needs to be aligned to a group template and resampled to a common grid
  - This allows for a voxelise comparisons across subjects

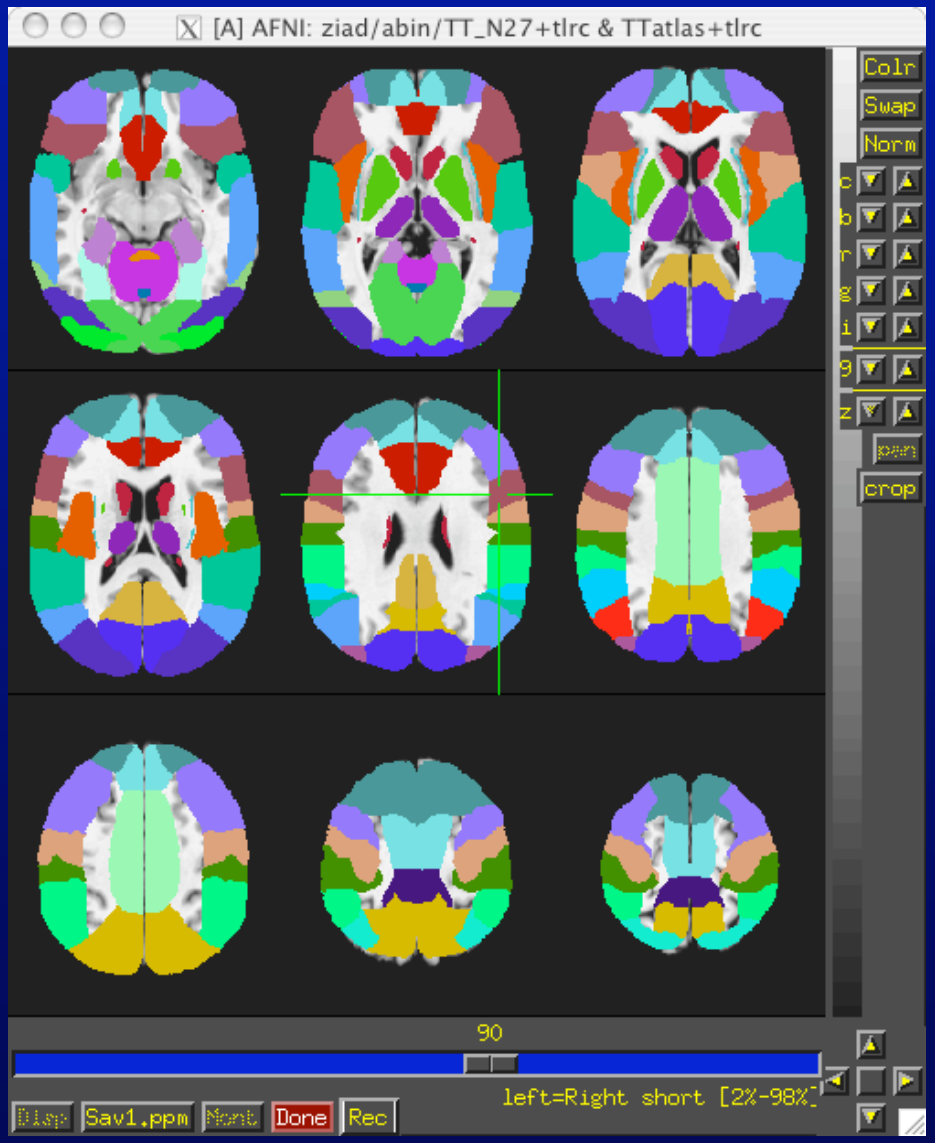


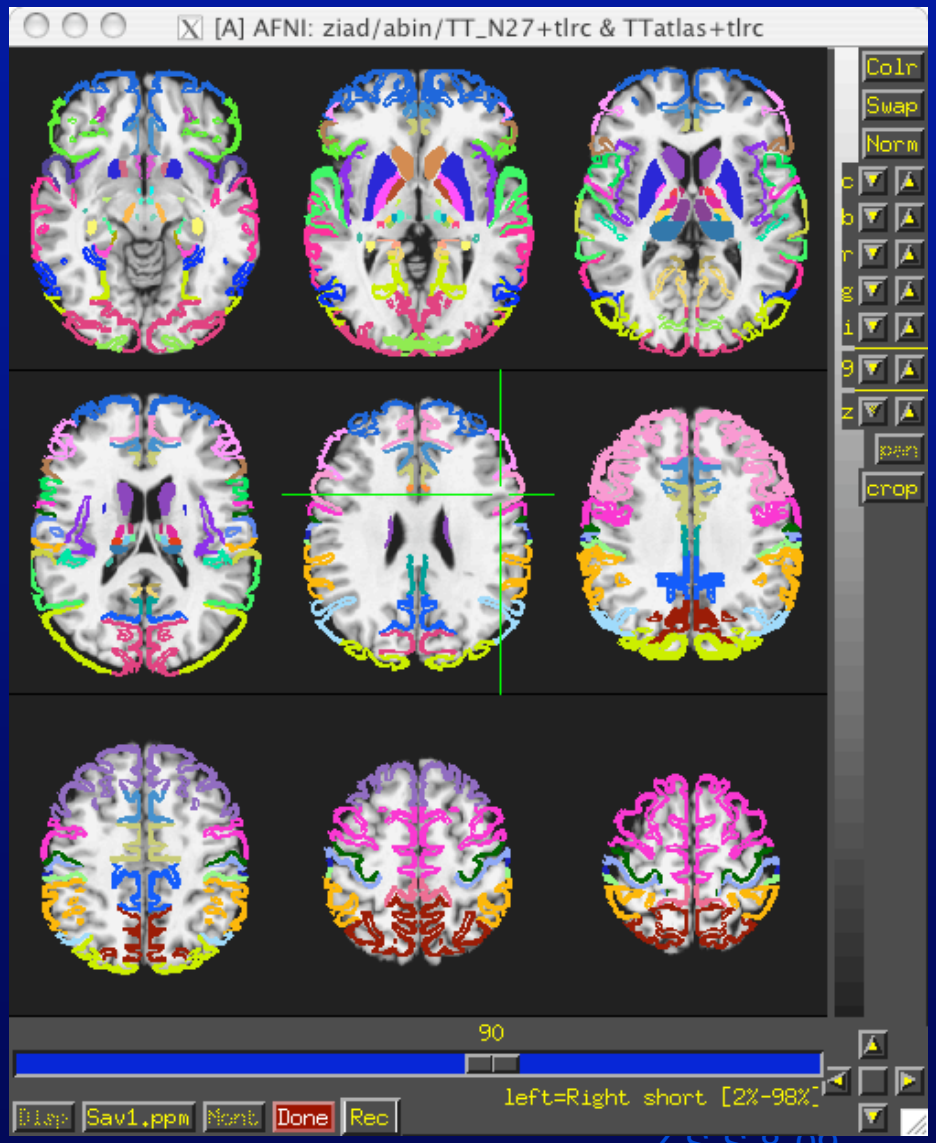
TLRC
MNI
MNI-Anat.

Atlases Distributed With AFNI

TT_Daemon

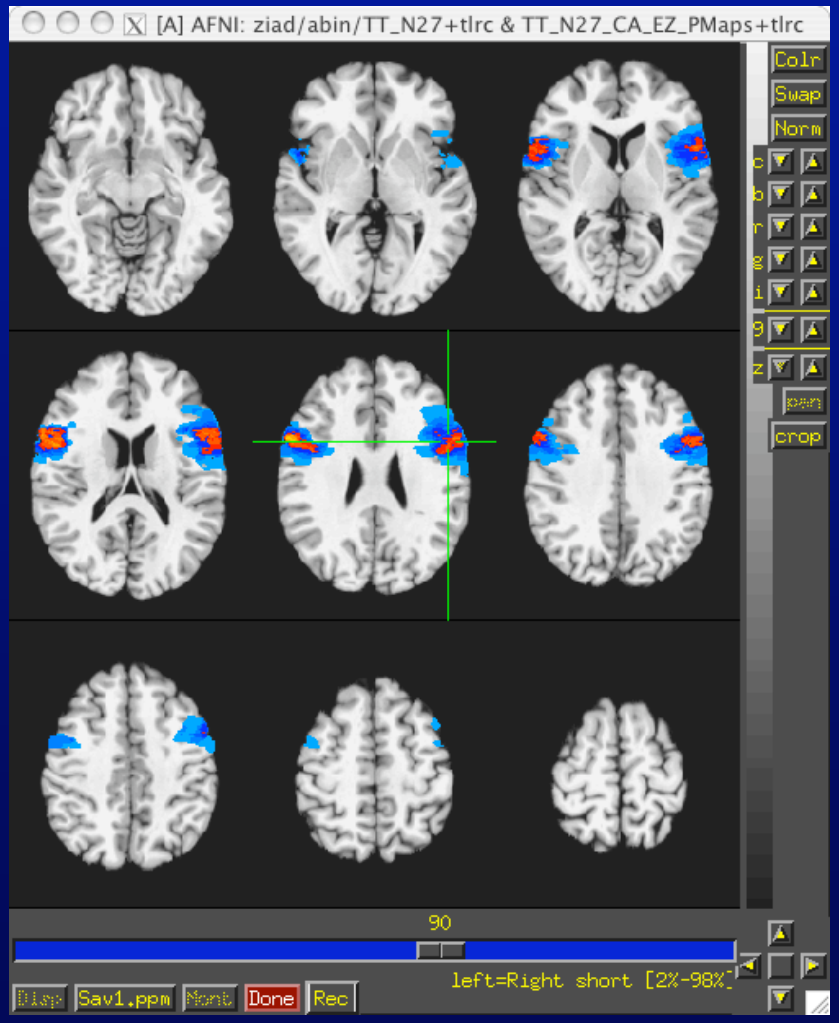
- TT_Daemon : Created by tracing Talairach and Tournoux brain illustrations.
  - Generously contributed by Jack Lancaster and Peter Fox of RIC UTHSCSA)

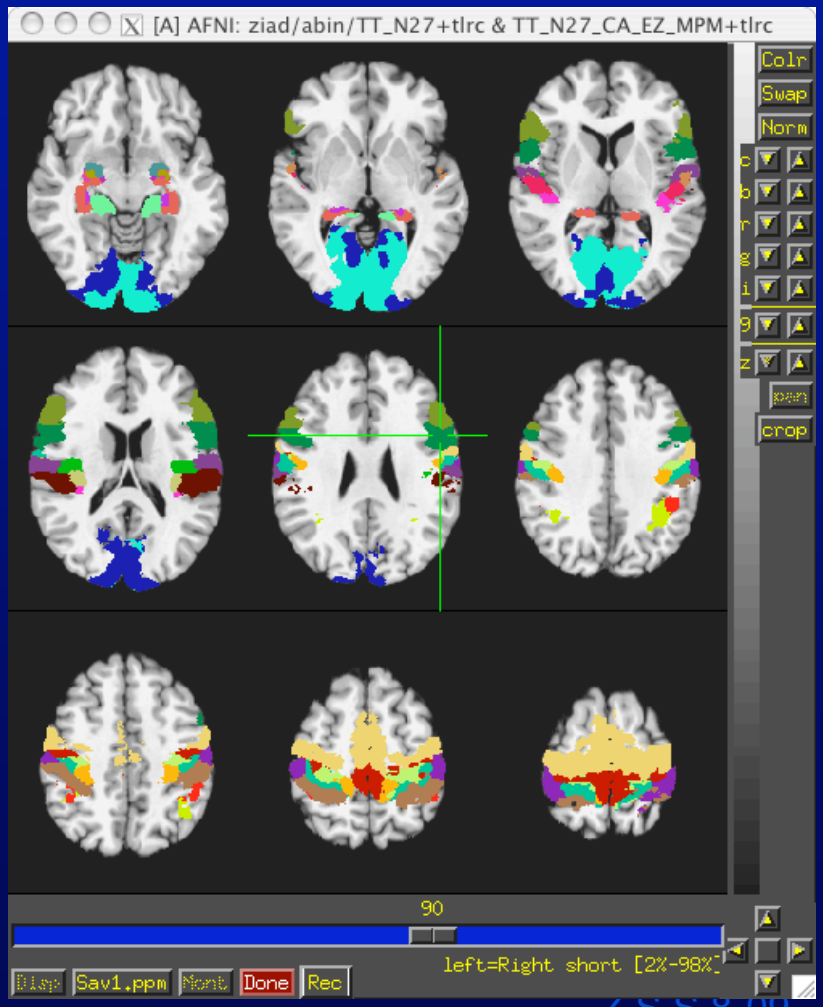




Atlases Distributed With AFNI
Anatomy Toolbox: Prob. Maps, Max. Prob. Maps

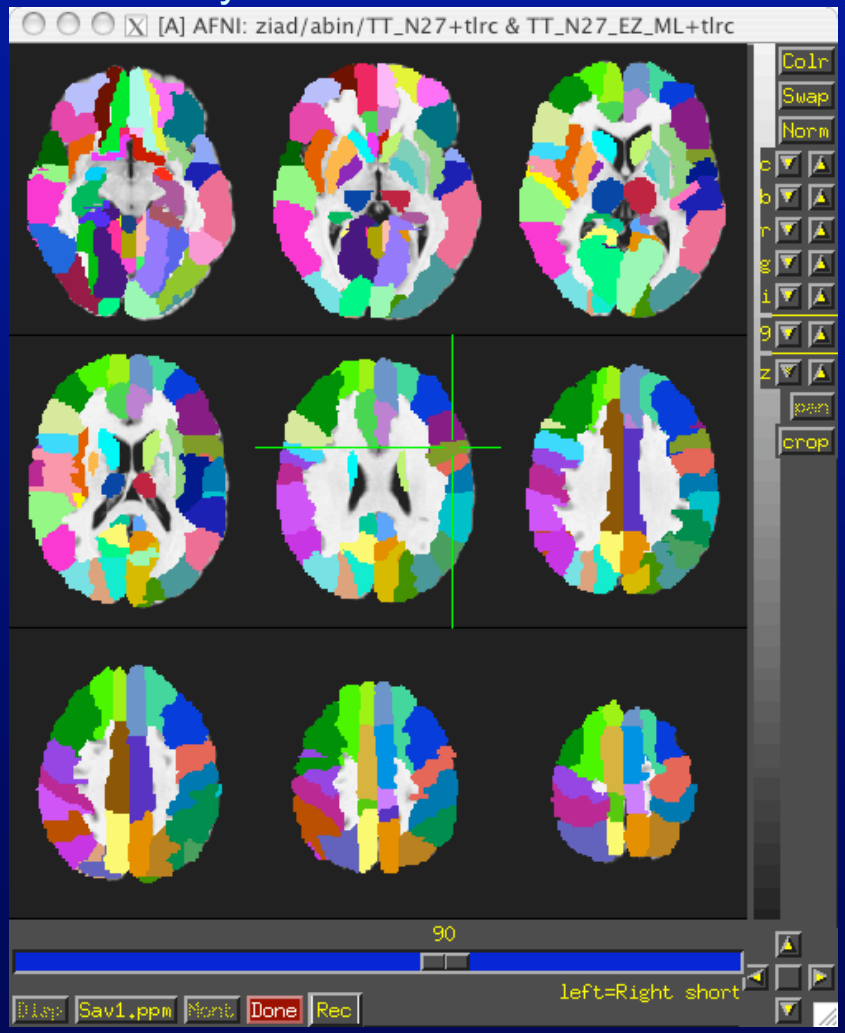
- CA_N27_MPM, CA_N27_ML, CA_N27_PM: Anatomy Toolbox's atlases with some created from cytoarchitectonic studies of 10 human post-mortem brains
  - Generously contributed by Simon Eickhoff, Katrin Amunts and Karl Zilles of IME, Julich, Germany **Eickhoff S. et al. 05**

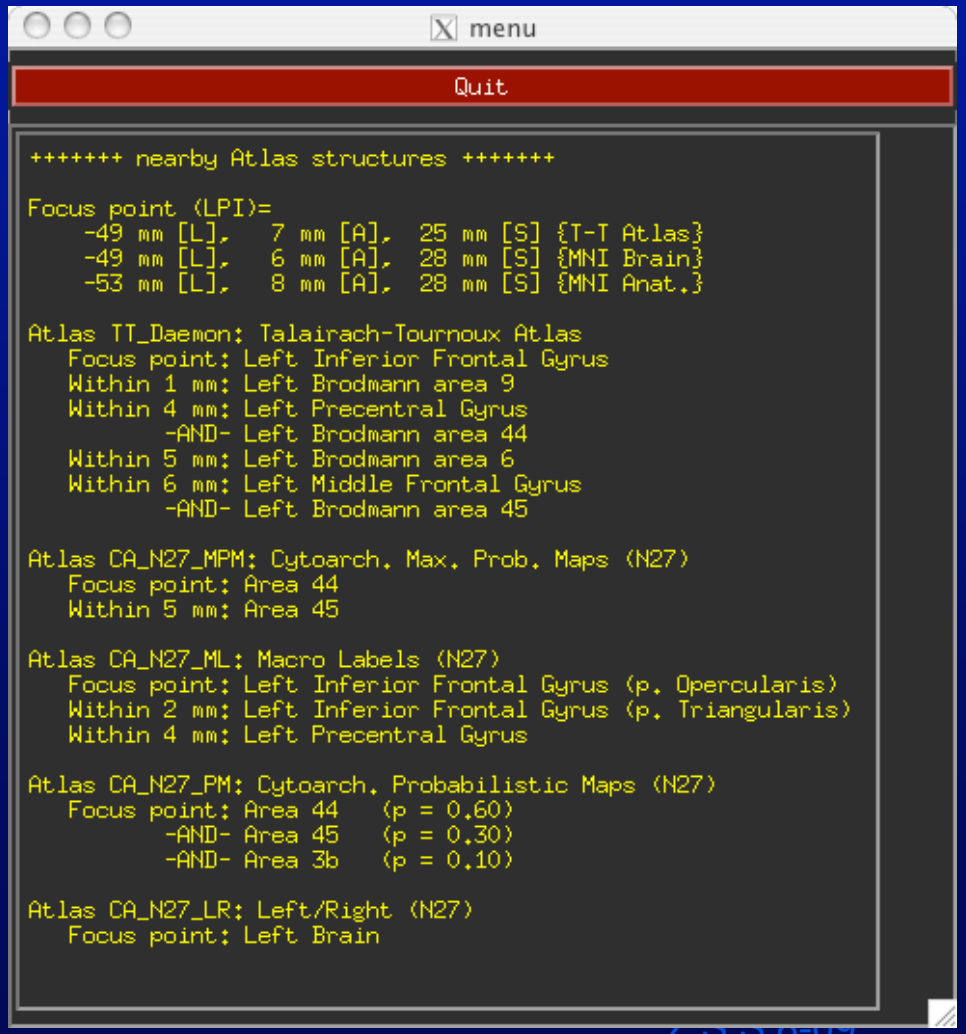




Atlases Distributed With AFNI: Anatomy Toolbox: MacroLabels

- CA_N27_MPM, CA_N27_ML, CA_N27_PM: Anatomy Toolbox's atlases with some created from cytoarchitectonic studies of 10 human post-mortem brains
  - Generously contributed by Simon Eickhoff, Katrin Amunts and Karl Zilles of IME, Julich, Germany





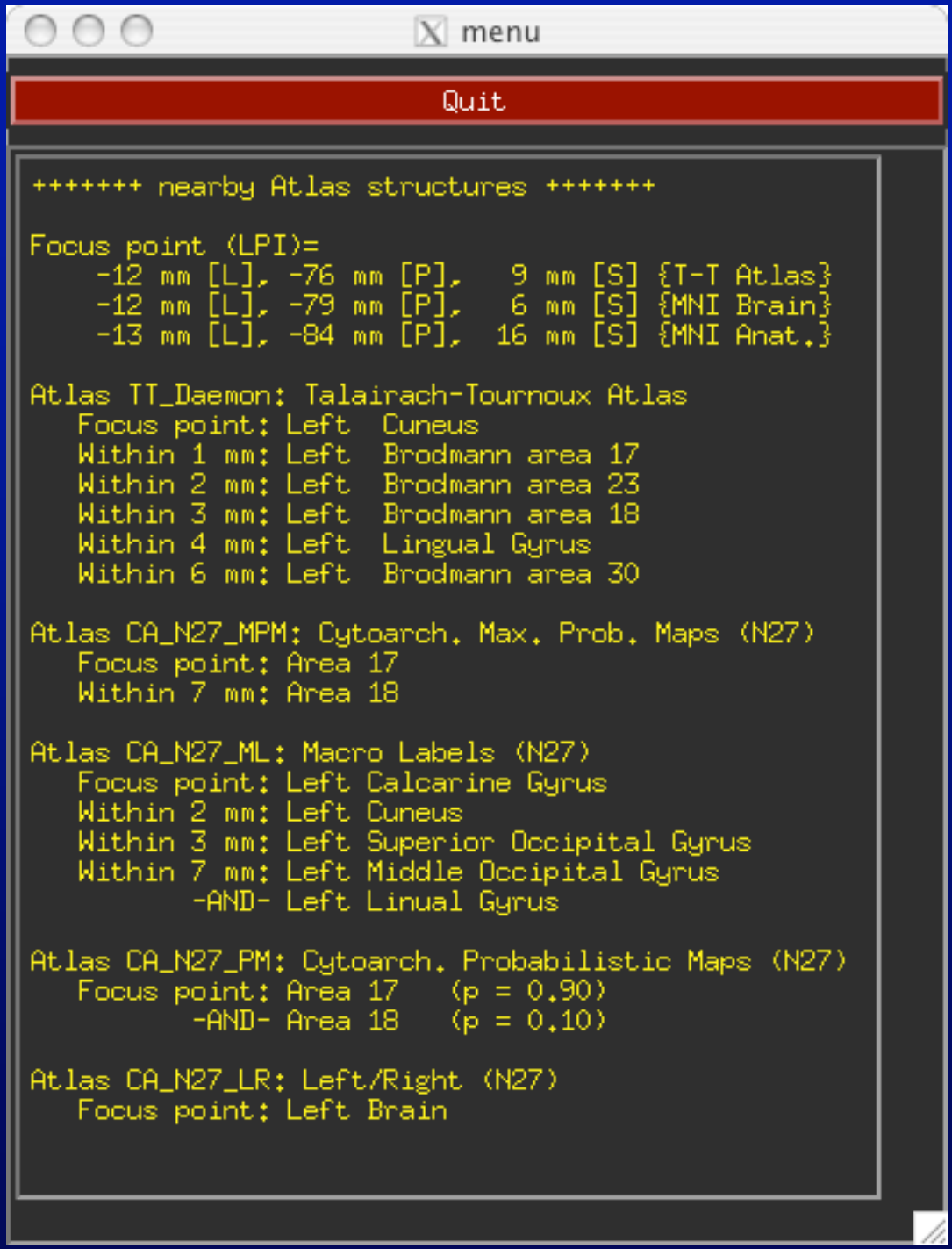
[Where am I?]

Shows you where you are in various atlases.

(works in `+orig` too, if you have a TT transformed parent)

For atlas installation, and much much more, see help in command line version:

`whereami -help`



- In this example, 4 ROI clusters were found that fit the criteria designated by the **`3dclust`** command. Below is an explanation of the output:

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
#Volume	CM RL	CM AP	CM IS	minRL	maxRL	minAP	maxAP	minIS	maxIS	Mean	SEM	Max Int	MI RL	MI AP	MI IS
#------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------	-------	-------	-----	-----	-----
387	-37.2	29.1	1.4	-61.4	-16.4	-1.3	58.7	-19.4	26.1	16.064	0.5	94.727	-38.9	28.7	-1.9
348	34.2	31.3	3.1	9.8	47.3	-1.3	58.7	-12.4	26.1	16.24	0.5369	80.252	32.3	28.7	1.6
75	-26.6	24.8	37.8	-42.7	-20.2	17.5	32.5	26.1	54.1	12.647	0.5967	29.788	-23.9	28.7	40.1
66	16.8	26.3	40.9	9.8	24.8	17.5	36.2	29.6	54.1	11.69	0.4156	22.931	17.3	21.2	40.1
#------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-------	-------	-------	-----	-----	-----
# 876	-3.7	29.6	6.9							15.512	0.3164				

  - Volume: Size of each cluster volume
  - CM RL: Center of mass (CM) for each cluster in the Right-Left direction
  - CM AP: Center of mass for each cluster in the Anterior-Posterior direction
  - CM IS: Center of mass for each cluster in the Inferior-Superior direction
  - minRL,maxRL: Bounding box for cluster, min & max coordinates in R-L direction
  - minAP,maxAP: Bounding box for cluster, min & max coordinates in A-P direction
  - minIS, maxIS: Bounding box for cluster, min & max coordinates in I-S direction
  - Mean: Mean value for each volume cluster
  - SEM: Standard error of the mean for the volume cluster
  - Max Int: Maximum Intensity value for each volume cluster
  - MI RL: Maximum Intensity value in the R-L direction of each volume cluster
  - MI AP: Maximum intensity value in the A-P direction of each volume cluster
  - MI IS: Maximum intensity value in the I-S direction of each volume cluster

- **whereami** can report on the overlap of ROIs with atlas-defined regions

whereami -omask anat_roi+tlrc

```
++ Input coordinates orientation set by default rules to RAI
++ Input coordinates space set by default rules to TLRC
++ In ordered mode ...
++ Have 2 unique values of:
   0   1
++ Skipping unique value of 0
++ Processing unique value of 1
++    195 voxels in ROI
++    195 voxels in atlas-resampled mask
Intersection of ROI (valued 1) with atlas TT_Daemon (sb0):
   89.2 % overlap with Middle Occipital Gyrus, code 33
   6.7  % overlap with Middle Temporal Gyrus, code 35
   -----
   95.9 % of cluster accounted for.

Intersection of ROI (valued 1) with atlas TT_Daemon (sb1):
   19.5 % overlap with Brodmann area 37, code 113
   1.5  % overlap with Brodmann area 19, code 96
   -----
   21.0 % of cluster accounted for.

++    195 voxels in atlas-resampled mask
Intersection of ROI (valued 1) with atlas CA_N27_MPM (sb0):
   1.5  % overlap with hOC5 (V5 / MT+), code 110
   -----
   1.5  % of cluster accounted for.

++    195 voxels in atlas-resampled mask
Intersection of ROI (valued 1) with atlas CA_N27_ML (sb0):
   61.0 % overlap with Right Middle Occipital Gyrus, code 52
   20.0 % overlap with Right Middle Temporal Gyrus, code 86
   -----
   81.0 % of cluster accounted for.
```

Localization In Standard Spaces

- Single location not enough
- Specify units, template, and space
- Describe coverage
- Look at the data, know the anatomy

Multi-Voxel Statistics

Spatial Clustering
&
False Discovery Rate:

"Correcting" the Significance

Basic Problem

- Usually have 20-100K FMRI voxels in the brain
- Have to make at least one decision about each one:
  - Is it “active”?
    - That is, does its time series match the temporal pattern of activity we expect?
  - Is it differentially active?
    - That is, is the BOLD signal change in task #1 different from task #2?
- Statistical analysis is designed to control the error rate of these decisions
  - Making ***lots*** of decisions: hard to get perfection in statistical testing

Multiple Testing Corrections

- **Two types of errors**
  - What is **H_0** in FMRI studies? H_0: no effect (activation, difference, …) at a voxel
  - Type I error = Prob(reject H_0 when H_0 is true) = false positive = p value

    Type II error = Prob(accept H_0 when H_1 is true) = false negative = β

    power = $1-\beta$ = probability of detecting true activation
  - Strategy: control type I error while increasing power (decrease type II errors)
  - Significance level α (magic number 0.05) : $p < \alpha$

Justice System: Trial

Hidden Truth

	Defendant Innocent	Defendant Guilty
Reject Presumption of Innocence (Guilty Verdict)	**Type I Error (defendant very unhappy)**	**Correct**
Fail to Reject Presumption of Innocence (Not Guilty Verdict)	**Correct**	**Type II Error (defendant very happy)**

Statistics: Hypothesis Test

Hidden Truth

	H_0 True Not Activated	H_0 False Activated
Reject H_0 (decide voxel is activated)	**Type I Error (false positive)**	**Correct**
Don't Reject H_0 (decide voxel isn't activated)	**Correct**	**Type II Error (false negative)**

- **Family-Wise Error (FWE)**
  - Simple probability example: sex ratio at birth = 1:1
    - Chance there are 5 boys in a family with 5 kids: $(1/2)^5 \approx 0.03$
    - For 10,000 families with 5 kids, expected #families with 5 boys: $10{,}000 \times (2)^{-5} \approx 312$
  - Multiple testing problem: voxel-wise statistical analysis
    - With N voxels, what is the chance to make a false positive error (Type I) in one or more voxels?

      Family-Wise Error: $\alpha_{FW} = 1-(1-p)^N \rightarrow 1$ as N increases
    - For $N \cdot p$ small (compared to 1), $\alpha_{FW} \approx N \cdot p$
    - $N \approx 20{,}000+$ voxels in the brain
    - To keep probability of even one false positive $\alpha_{FW} < 0.05$ (the "corrected" p-value), need to have $p < 0.05 / 2\times10^4 = 2.5\times10^{-6}$
    - This constraint on the per-voxel ("uncorrected") p-value is so stringent that we'll end up rejecting a lot of true positives (Type II errors) also, just to be safe on the Type I error rate
  - Group analysis is the most severe situation
    - (have the least data, considered as number of independent samples = subjects)

Approaches to the "Curse of Multiple Comparisons"

- Bonferroni correction
  - Use $p = \alpha / N$ as individual voxel significance level to achieve $\alpha_{FW} = \alpha$
    - Too stringent and overly conservative: $p = 10^{-8} \ldots 10^{-6}$
- Rescue from hell of statistical super-conservatism?
  - Correlation: Voxels in the brain are not independent
    - Especially after we smooth them together!
    - Means that Bonferroni correction is *way way* too stringent
  - Contiguity: Structures in the brain activation map
    - We are looking for activated "blobs": the chance that pure noise (H_0) will give a set of seemingly-activated voxels next to each other is lower than getting false positives that are scattered apart
  - Control FWE based on spatial correlation **and** minimum cluster size we are willing to accept
- Control false discovery rate (FDR)
  - FDR = expected proportion of false positive voxels among all detected voxels

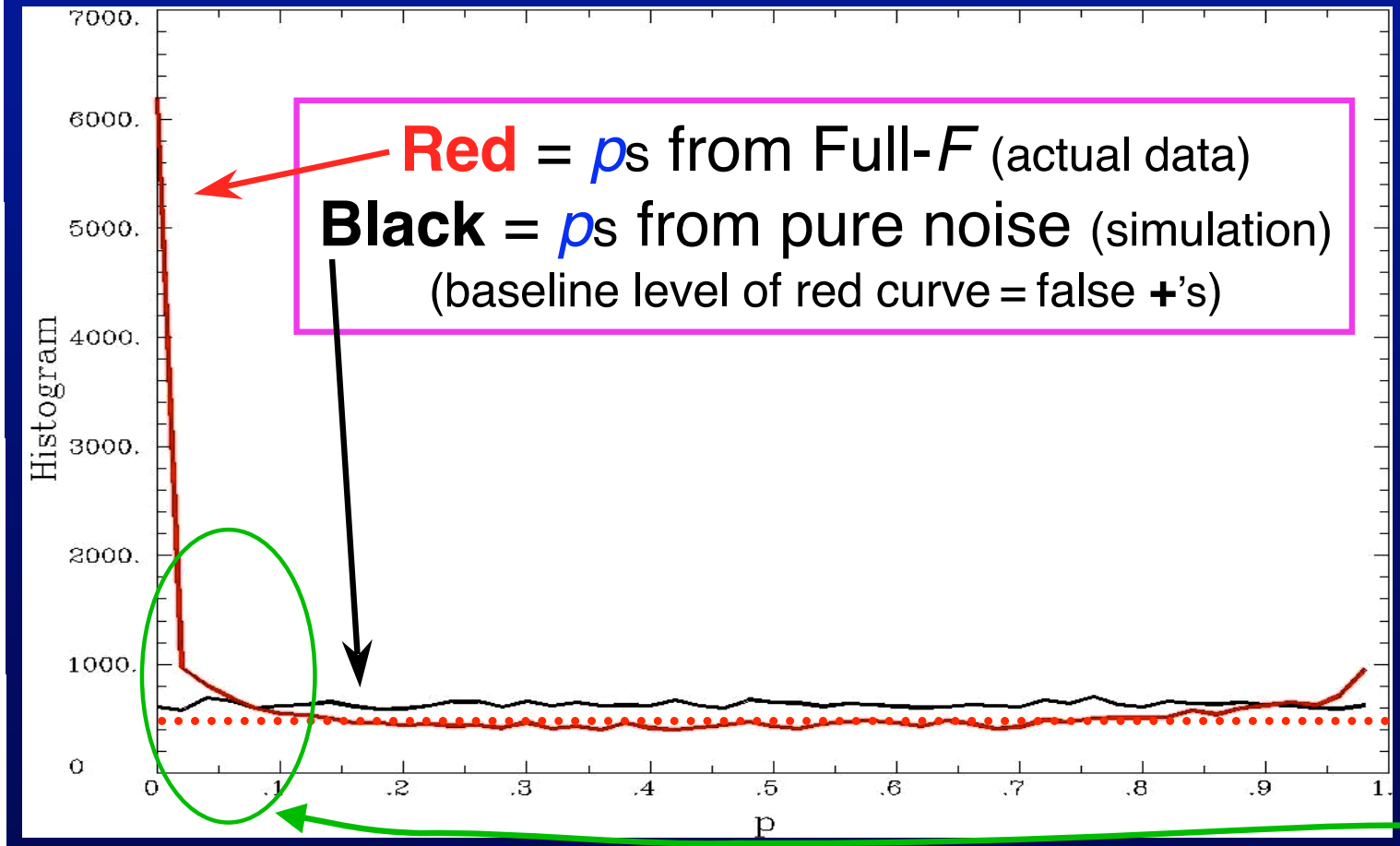
- **Example:** `AlphaSim -nxyz 64 64 20 -dxyz 3 3 5 \`
  `-fwhm 5 -pthr 0.001 -iter 1000`
- Output is in 6 columns: focus on 1st and 6th columns (ignore others)
  - `Cl Size  Frequency  CumuProp  Alpha`

1	47064	0.751113	1.000000
2	11161	0.929236	1.000000
…	…	…	…
6	111	0.998995	0.158000
7	32	0.999505	0.058000
8	20	0.999825	0.029000
9	8	0.999952	0.010000
10	2	0.999984	0.003000

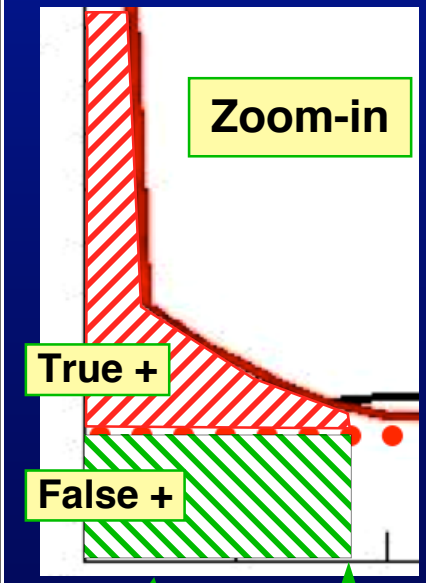
  - At this uncorrected $p = 0.001$, in this size volume, with noise of this smoothness: the chance of a cluster of size 8 *or larger* occurring by chance alone is 0.029
  - May have to run several times with different uncorrected *p*
    - uncorrected *p* ↑ ⇔ required minimum cluster size ↑

Basic Ideas Behind FDR *q*

- ***If*** all the null hypotheses are true, ***then*** the statistical distribution of the *p*-values will be uniform
  - Deviations from uniformity at low *p*-values ⇒ true positives
  - Baseline of uniformity indicates how many true negatives are hidden amongst in the low *p*-value region



31,555 voxels
50 histogram bins



threshold at this *p*

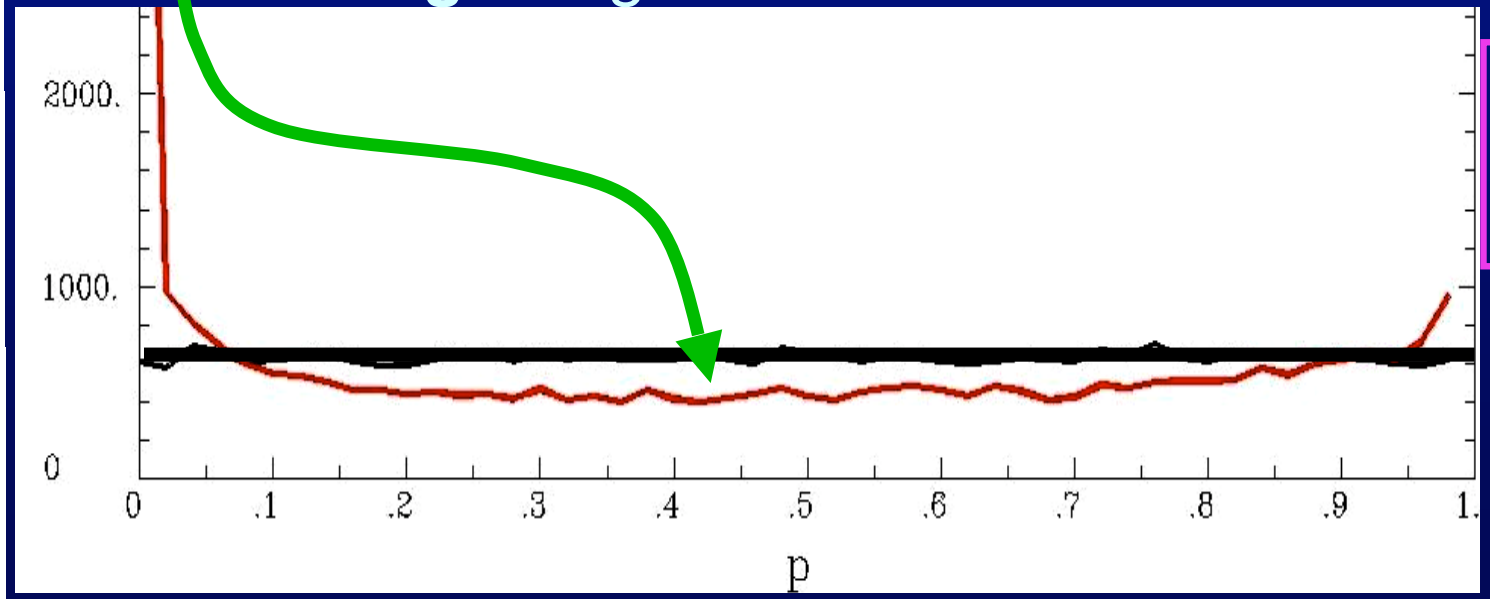
Z.S.S 8-09

Adapting FDR to get MDF - 1

- **MDF** = **M**issed **D**etection **F**raction
  - = fraction of true positives with *p*-value below a threshold
    - Raise threshold enough: MDF goes to 100% as FDR goes to 0%
  - = typically 50+% in FMRI: we're missing 1/2 the activations*!*

① Estimate m_1 = total fraction of true positives in data

- Compare bulk of *p*-value histogram with uniform distribution assuming no true positives
- Deficit between **data's *p* histogram** and **uniform histogram** gives estimated number of true positives



31,555 voxels
50 histogram bins
[Same data as before]

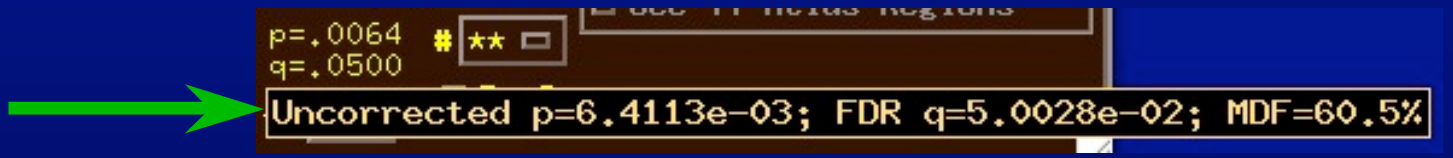
Adapting FDR to get MDF - 2

② At given threshold: have J detections out of N voxels

- # false detections ≈ qJ
- # true detections ≈ $(1-q)J$
- Fraction of true detections =

  # true detections ÷ # true positives ≈ $(1-q)J \div m_1 N$
- ∴ Missed detection fraction = MDF ≈ $1-[(1-q)J \div m_1 N]$

- MDF estimate is in a popup "hint" in AFNI GUI



- The key to getting MDF is a good estimator for m_1
  - Which is hard to do accurately (e.g., lots of assumptions)
  - So MDF is just a crude approximation at this time
  - Estimate of m_1 is also used to adjust FDR: $q' = (1-m_1)q$

Group (Stage 2) analysis

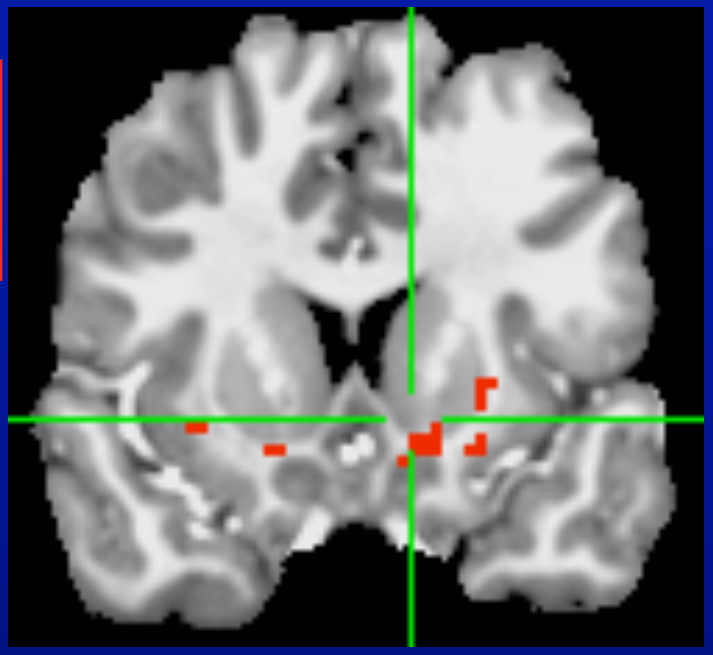
- Earlier approaches only carry beta coefficients to the group level analysis
  - Within/intra-subject variability (standard error, sampling error) is relatively small compared to cross/between/inter-subjects variability
  - Within/intra-subject variability roughly the same across subjects
    - TTest (paired, unpaired)
    - ANOVA (1-5 way)
    - 3dLME combination of random and fixed effects analysis
      - Unbalanced designs (unequal # of subjects, missing data, etc.)
      - ANOVA and ANCOVA, with unlimited # of factors & covariates
      - Violations of sphericity: heteroscedasticity, variance-covariance structure of observations (e.g., temporal correlation in HRF βs)

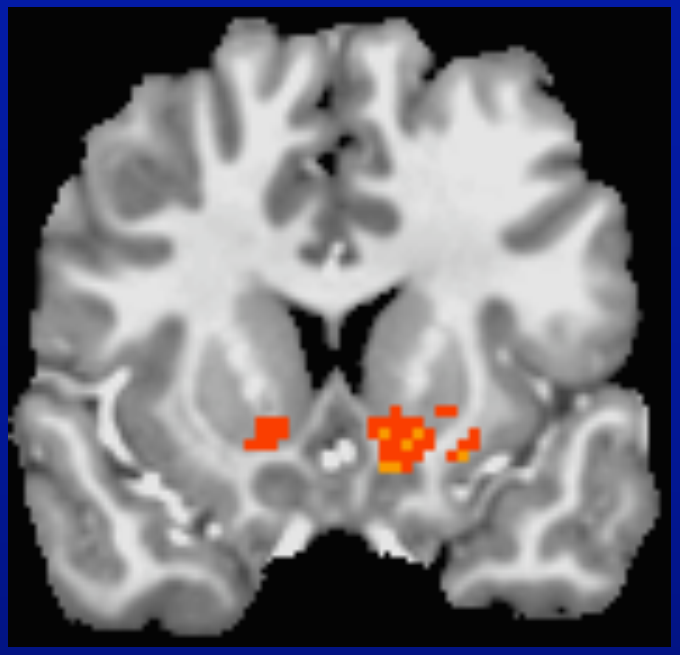
Group (Stage 2) analysis

- Newer approaches carry beta and variance
  - 3dMEMA (heteroscedasticity, account for beta variance)
    - Trust those β's with high reliability/precision (small SE or σ) through weighting
      - β with lower SE has more say in the final result
      - β with less significance gets downgraded
  - Results are more robust than earlier approaches
  - But more limited in types of tests than other such as 3dLME

Results: 3dANOVA vs. 3dMEMA

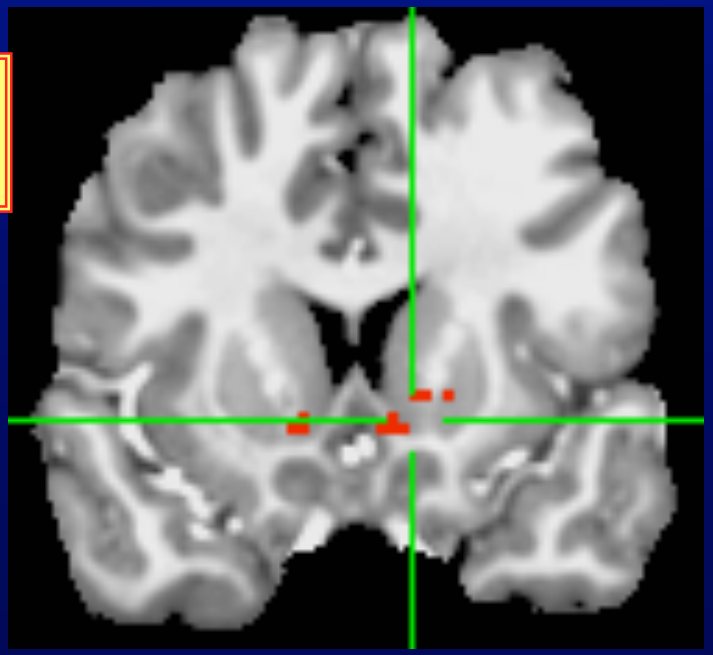
ANOVA: 12 Control subjects





MEMA: Control subjects

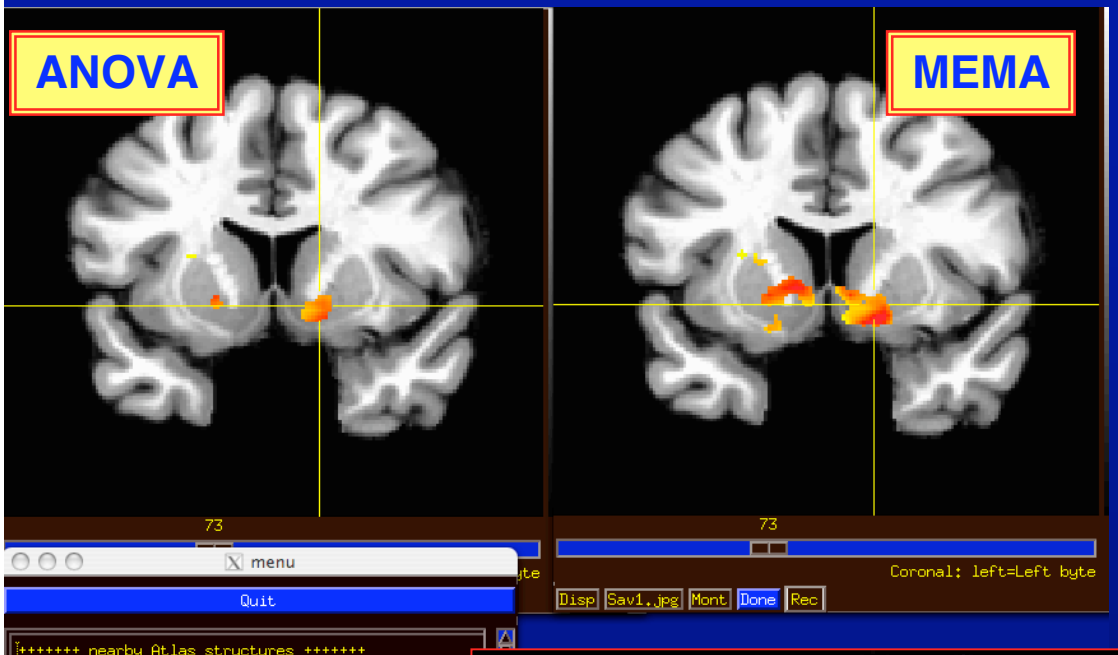
ANOVA: 12 Patients



MEMA: Patients

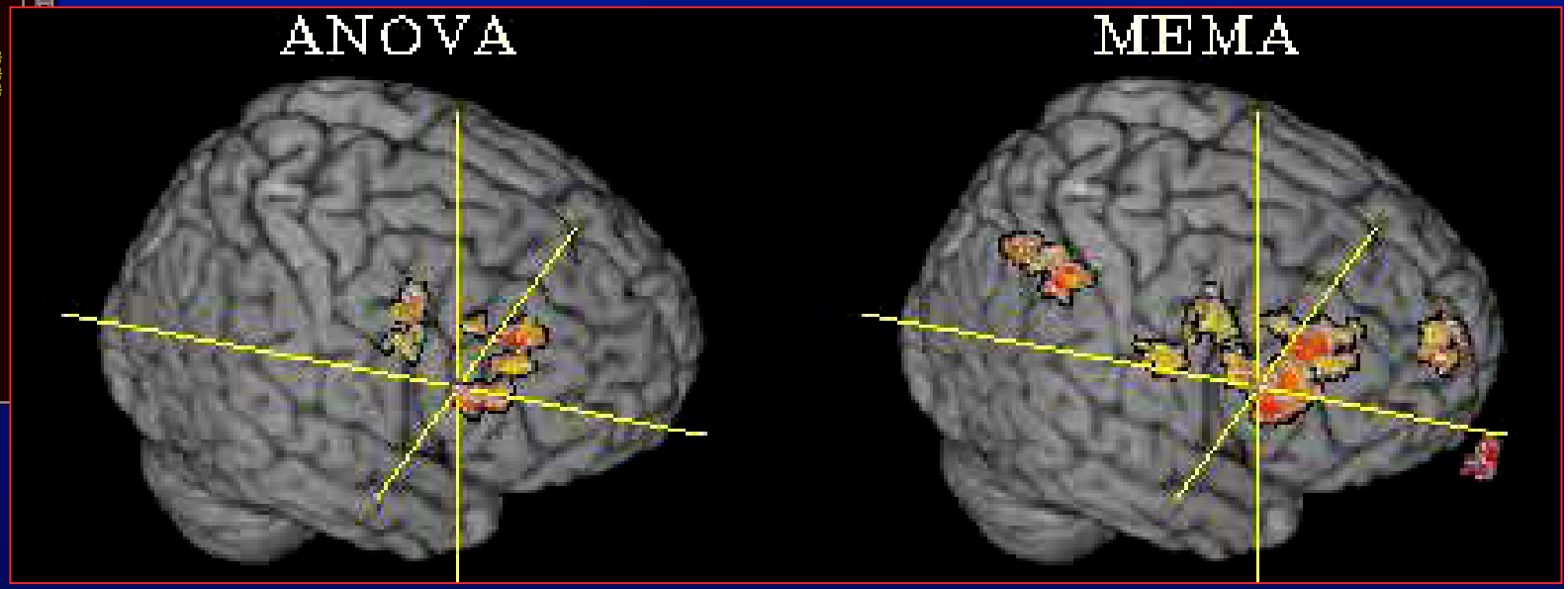
Data courtesy James Bjork NIDA/NIH

Results: 3dANOVA vs. 3dMEMA



Same uncorrected *p*-values;
ANOVA does not survive FDR;
MEMA laughs at such quibbles

Volume rendering movie pair made in AFNI : 5-10 min work

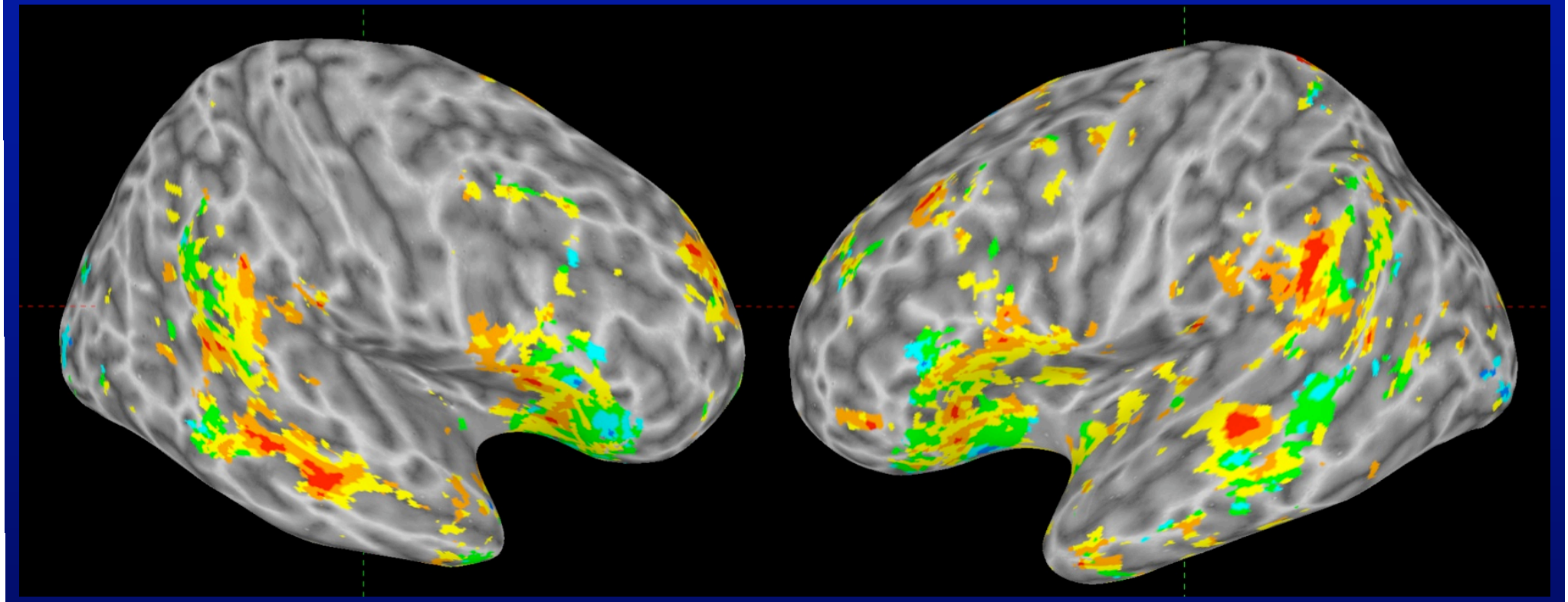


Data courtesy
James Bjork
NIDA/NIH

ventral striatal activation in an incentive task

Results: 3dttest vs. 3dMEMA

- Color = Difference between *t*-statistics in voxels with *t*(30) > 2



(red >= 2.8, 1.7 <= orange < 2.8; 0.5 <= yellow < 1.7; -0.5 <= green < 0.5; blue <= -0.5)

Majority of significant voxels gained power (warm colors)

Data courtesy
Vince Costa
U Florida

Start simple: one-sample test

- Random-effects: $y_i=\theta_i+\varepsilon_i=\alpha_0+\delta_i+\varepsilon_i$, for ith subject
  - y_i : β or linear combination (contrast) of β's from ith subject
  - $\theta_i=\alpha_0+\delta_i$: "true" individual effect from ith subject
  - α_0: group effect we'd like to find out
  - δ_i : deviation of ith subject from group effect α_0, $N(0, \tau^2)$
  - ε_i: sample error from ith subject, $N(0, \sigma_i^2)$, σ_i^2 known!
- Special cases
  - $\sigma_i^2=0$ reduced to conventional group analysis:

    One-sample t: $y_i = \alpha_0 + \delta_i$
  - $\delta_i=0$ ($\tau^2=0$) assumed in fixed-effects model: Ideally we could find out all possible explanatory variables so only an FE model is necessary!

MEMA with one-sample test

- Random-effects: $y_i = \alpha_0 + \delta_i + \varepsilon_i$, for ith subject
  - $\delta_i \sim N(0, \tau^2)$, $\varepsilon_i \sim N(0, \sigma_i^2)$, σ_i^2 known, τ^2 unknown
  - What can we achieve?
    - Null hypothesis about group effect H_0: $\alpha_0 = 0$
    - Checking group heterogeneity H_0: $\tau^2 = 0$
    - Any outliers among the subjects? Adding some confounding variable(s)? Grouping subjects?
  - We know σ_i^2, and pretend we also knew τ^2, weighted least squares (WLS) gives
    - The "best" estimate
    - **BLUE**: unbiased with minimum variance

$$\alpha_0 = \frac{\sum w_i y_i}{\sum w_i}, w_i = \frac{1}{\tau^2 + \sigma_i^2}$$

  - Unfortunately we don't know τ^2

Solving MEMA

- Estimating τ^2: a few approaches
  - Method of moment (MoM) - DSL
  - Maximum likelihood (ML)
  - Restricted/residual/reduced/marginal ML (REML): 3dMEMA
- Statistical testing
  - Group effect α_0=0:

$$Z = \frac{\sum w_i y_i}{\sqrt{\sum w_i}} \cong N(0,1), w_i = \frac{1}{\tau^2 + \sigma_i^2}$$

    - ❑Wald or *Z*-test: assume enough subjects with normal distributions
    - ❑Go with *t*-test when in doubt

$$Q = \sum_{i=1}^{n} \frac{(y_i - \alpha_0)^2}{\sigma_i^2} \sim \chi^2(n-1)$$

  - Heterogeneity test τ^2=0:
  - Outlier identification for each subject through *Z*-statistic

A slightly more complicated case

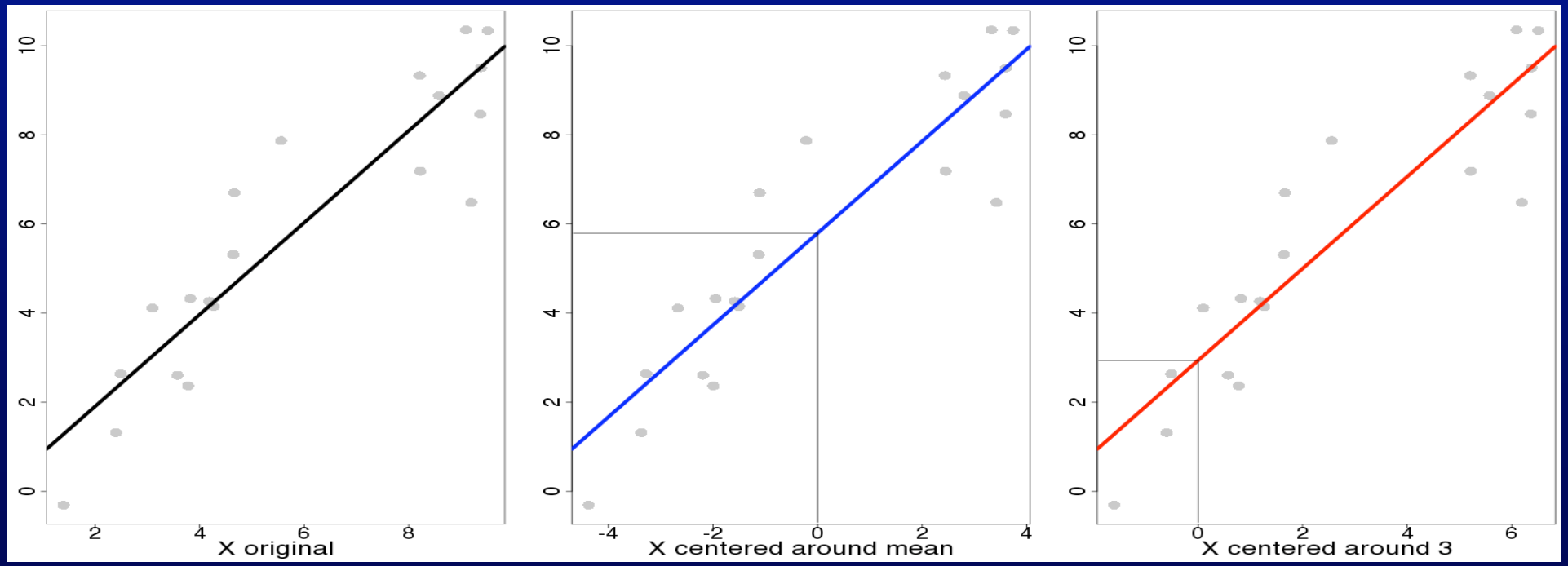
- $y_i = \alpha_0 + \alpha_1 x_{i1} + \ldots + \alpha_{ip} x_{ip} + \delta_i + \varepsilon_i$, for ith subject
  - Mixed-effects model or meta regression
  - y_i: β or linear combination (contrast) of β's from ith subject
  - α_0: common group effect we'd like to find out
  - x_{ij}: an indicator/dummy variable showing, for example, group to which ith subject belongs, level at which a factor lies, or a continuous variable such as covariate (e.g., age, IQ) (j=1,…,p)
  - δ_i : deviation of ith subject from group effect α_0, $N(0, \tau^2)$
  - ε_i: sample error from ith subject, $N(0, \sigma_i^2)$, σ_i^2 known!
- Combine subjects into a concise model in matrix form
  - $\mathbf{y}_{n\times 1} = \mathbf{X}_{n\times p}\alpha_{p\times 1} + \delta_{n\times 1} + \varepsilon_{n\times 1}$
  - $\mathbf{y} \sim N(\mathbf{X}\alpha, \tau^2\mathbf{I}_n + \mathbf{V})$, $\mathbf{V} = \mathrm{diag}(\sigma_1, \ldots, \sigma_n)$ known, τ^2 unknown
  - Estimate α and τ^2 simultaneously via maximizing REML

Covariates

- Covariates
  - May or may not be of direct interest
  - Confounding, nuisance, or interacting variables
  - Subject-level
  - Continuous or discrete
  - One-sample model $y_i = \alpha_0+\alpha_1 x_i+\delta_i + \varepsilon_i$, for *i*th subject
  - Two-sample model $y_i = \alpha_0+\alpha_1 x_{1i}+\alpha_2 x_{2i}+\alpha_3 x_{3i}+\delta_i + \varepsilon_i$
- Examples
  - Age, IQ, brain volume, cortex thickness
  - Behavioral data

Handling covariates: one group

- Centering:
  - $y_i = \alpha_0+\alpha_1 x_i+\delta_i + \varepsilon$, for *i*th subject
  - Interested in group effect α_0 (x=0) while controlling (partialling out) x
  - α_1 - slope (change rate): % signal change per unit of x
  - Interpretability: group effect α_0 at what value of x: mean or any other value?

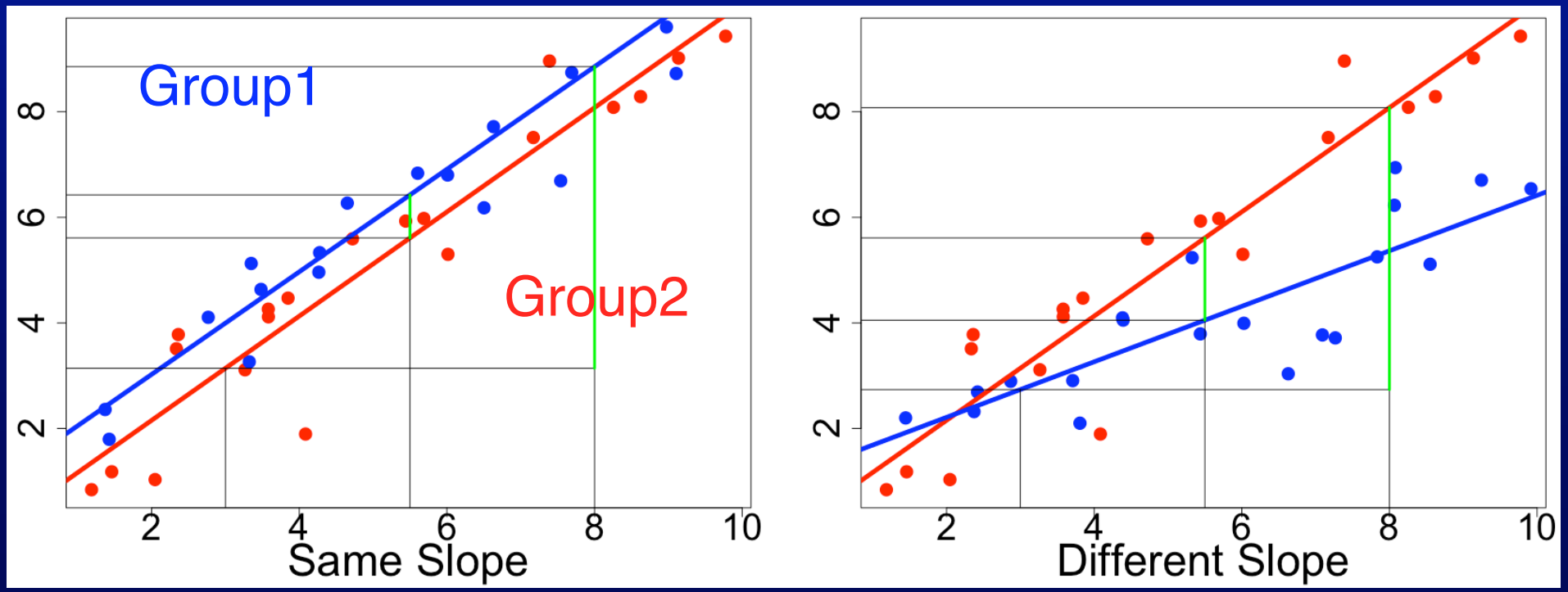


Covariates: trickier with > 1 group

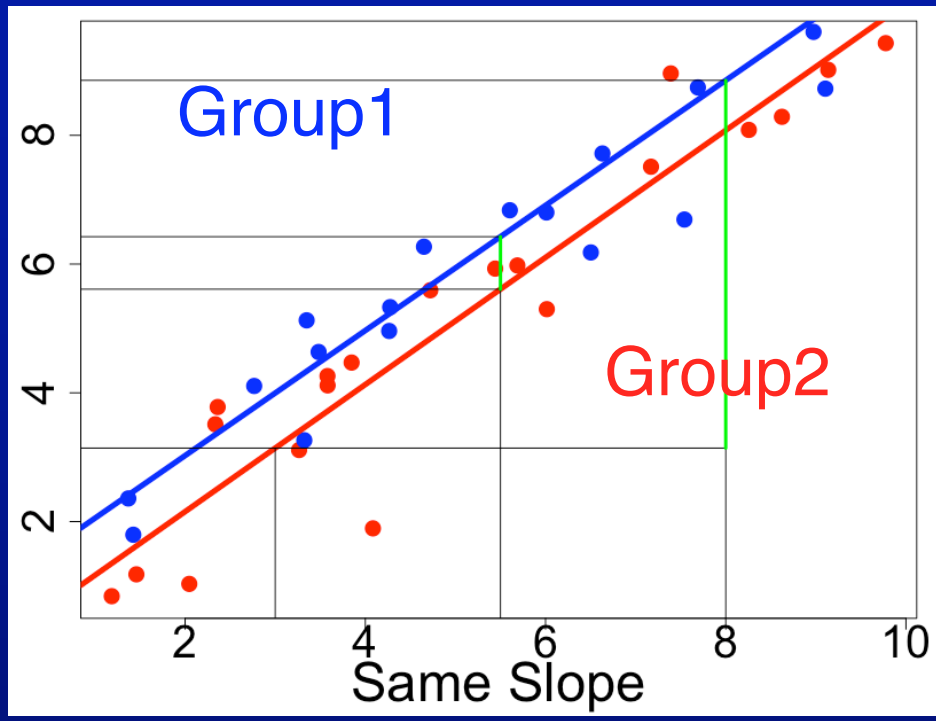
- Center and slope
  - $y_i = \alpha_0+\alpha_1 x_{1i}+\alpha_2 x_{2i}+\alpha_3 x_{3i}+\delta_i + \varepsilon$, for ith subject
    - x_1: group indicator
    - x_2: covariate
    - x_3: group effect in slope (interaction btw group and covariate)
  - What we're interested in
    - Group effects α_0 and α_1 while controlling covariate
  - Interpretability
    - Center
      - Group effect α_0 and α_1 at what covariate value?
      - Same or different center across groups?
    - Slope
      - same ($\alpha_3=0$) or different ($\alpha_3 \neq 0$) slope across groups

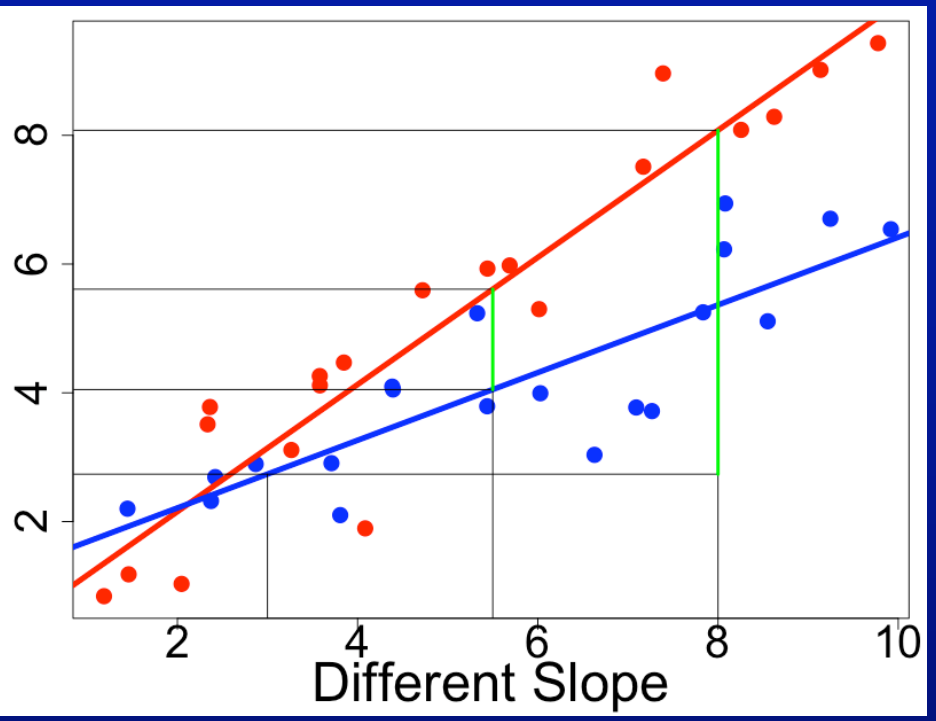
Covariates: scenarios with 2 groups

- Center and slope
  - $y_i = \alpha_0+\alpha_1 x_{1i}+\alpha_2 x_{2i}+\alpha_3 x_{3i}+\delta_i + \varepsilon_i$, for ith subject
  - Interpretability
    - Same center and same slope ($\alpha_3=0$)
    - Different center with same slope ($\alpha_3=0$)
    - Same center with different slope ($\alpha_3\neq 0$)
    - Different center and different slope ($\alpha_3\neq 0$)



Covariates: scenarios with 2 groups

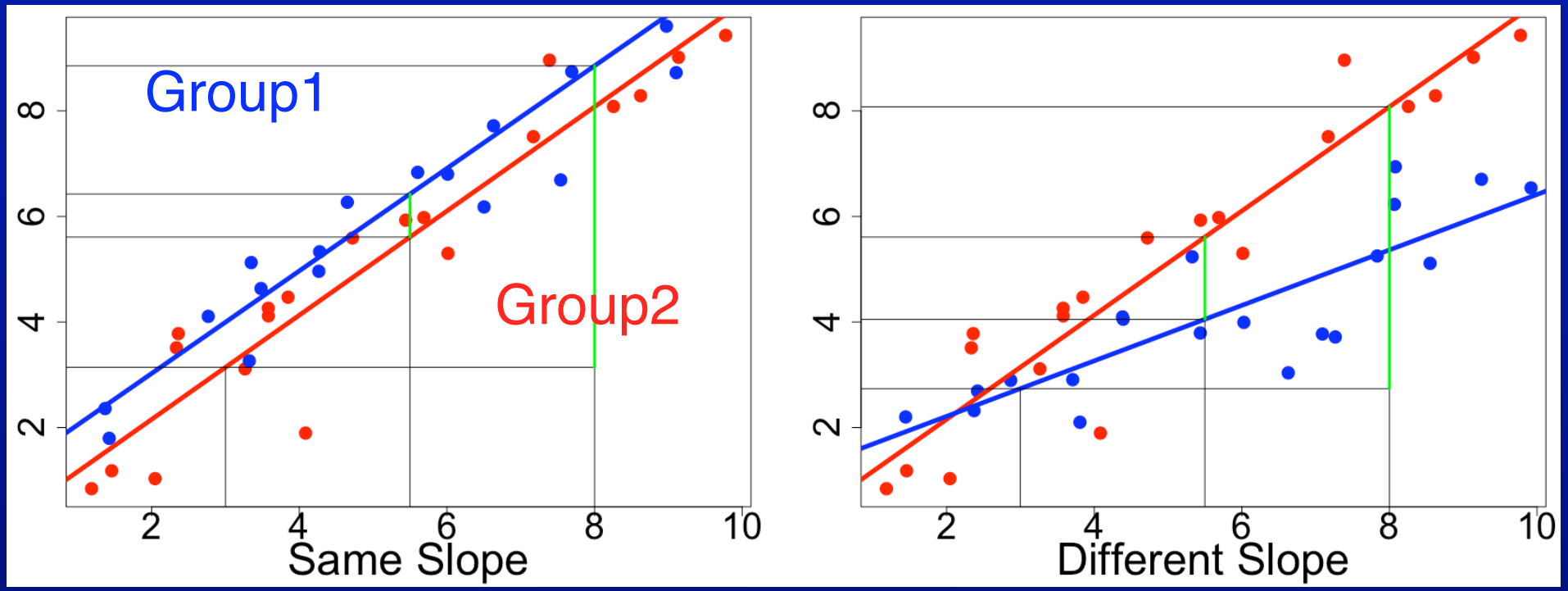




- If covariate for each group is centered on same value
  - Group effect constant regardless of that value
- Else group effect depends on centering difference

- When slopes are different:
  - Group effect depends on covariate value, even when centering is properly done

Covariates: scenarios with 2 groups



- Just "Regressing Out" a covariate is not enough!
  - Need to center properly
  - Need to consider covariate value (if α_3 significant)

Notes on a scandal

- Beware selection bias / circularity

Circular analysis in systems neuroscience: the dangers of double dipping.
Kriegeskorte et al. Nature Neuroscience 12(5) 2009

Voodoo correlations in social neuroscience
(*now known as* Puzzlingly High Correlations in fMRI Studies of Emotion, Personality, and Social Cognition)
Vul et al. Perspectives in Psychological Science 2009
Correlations in Social Neuroscience Aren't Voodoo: Commentary on Vul et al.
Lieberman et al. Perspectives in Psychological Science 2009

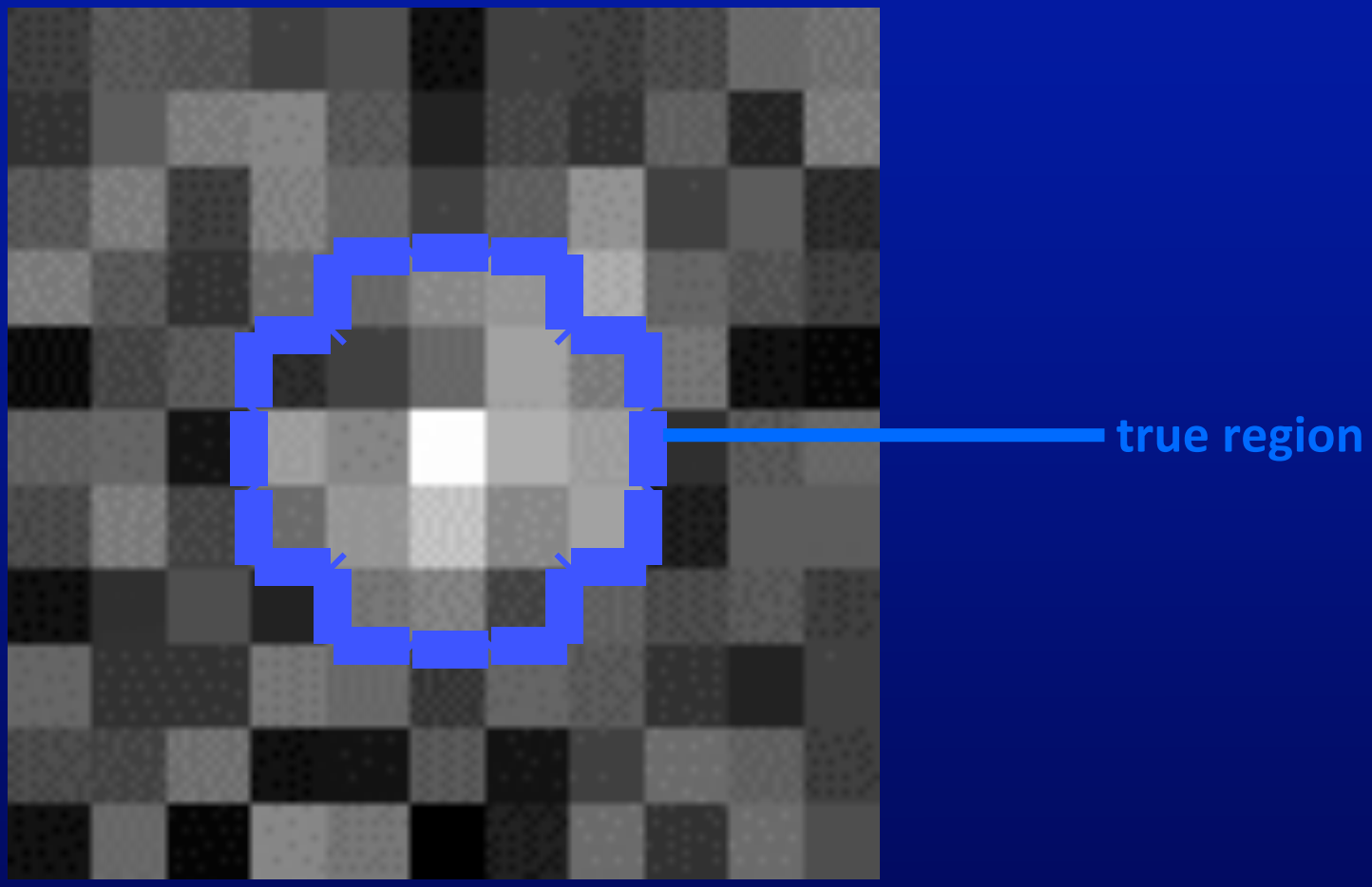
… (6 more commentaries/responses!) …

Big Correlations in Little Studies Inflated fMRI Correlations Reflect Low Statistical Power
Tal Yarkoni Perspectives in Psychological Science 2009

- Avoid hysteria
- Consider your false negatives

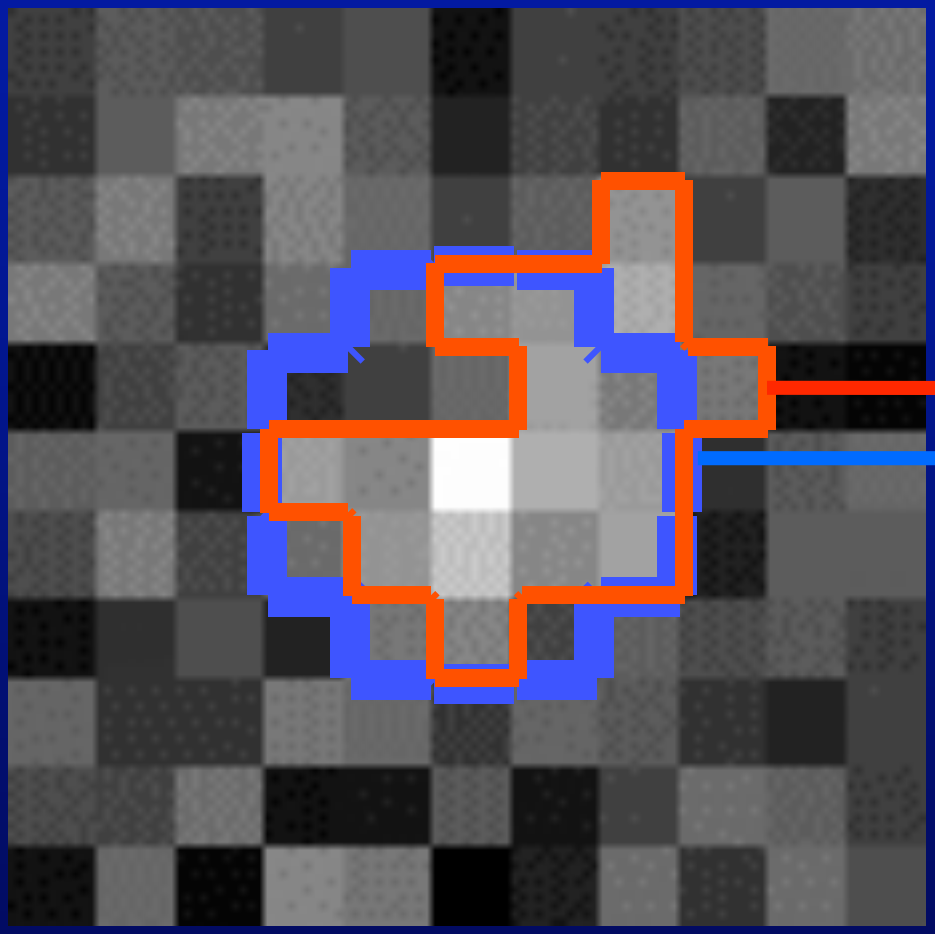
Z.S.S 8-09

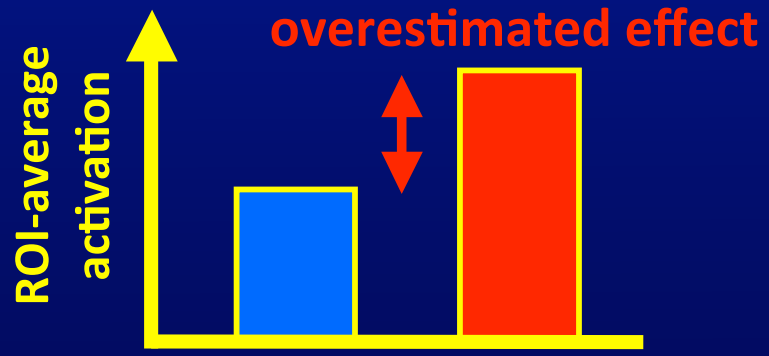
ROI selection bias



Slide from N. Kriegeskorte

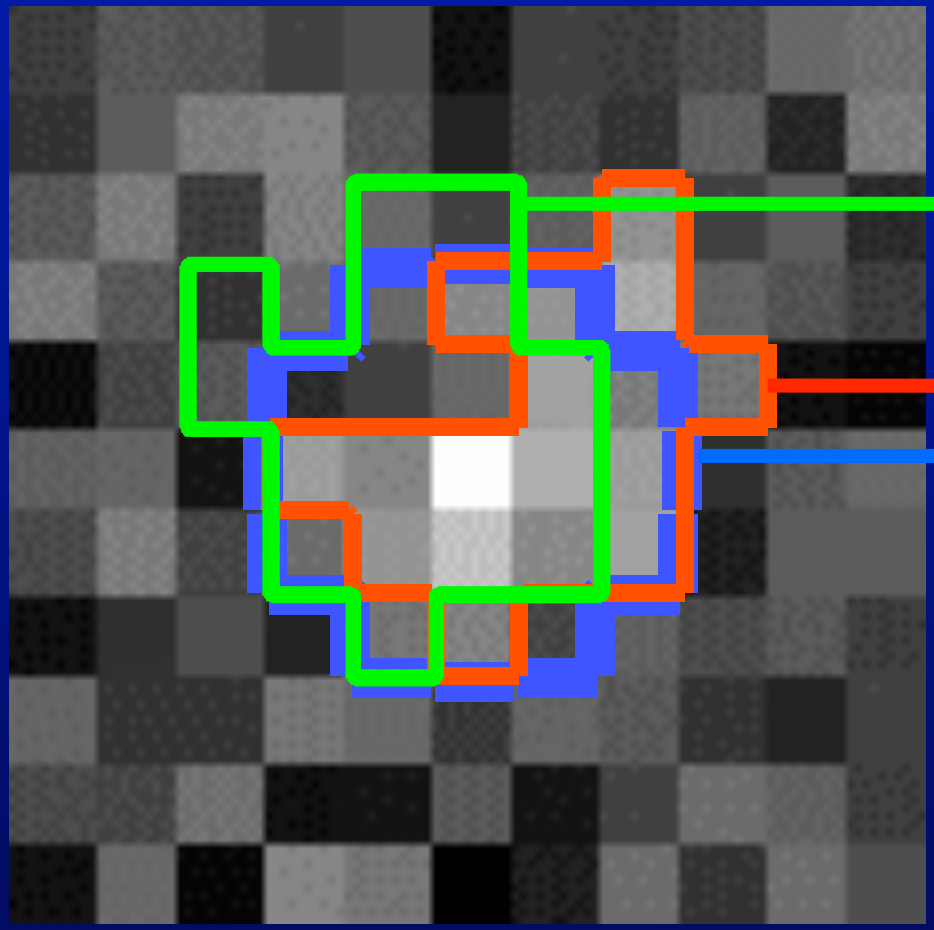
ROI selection bias





Slide from N. Kriegeskorte

ROI definition is affected by noise



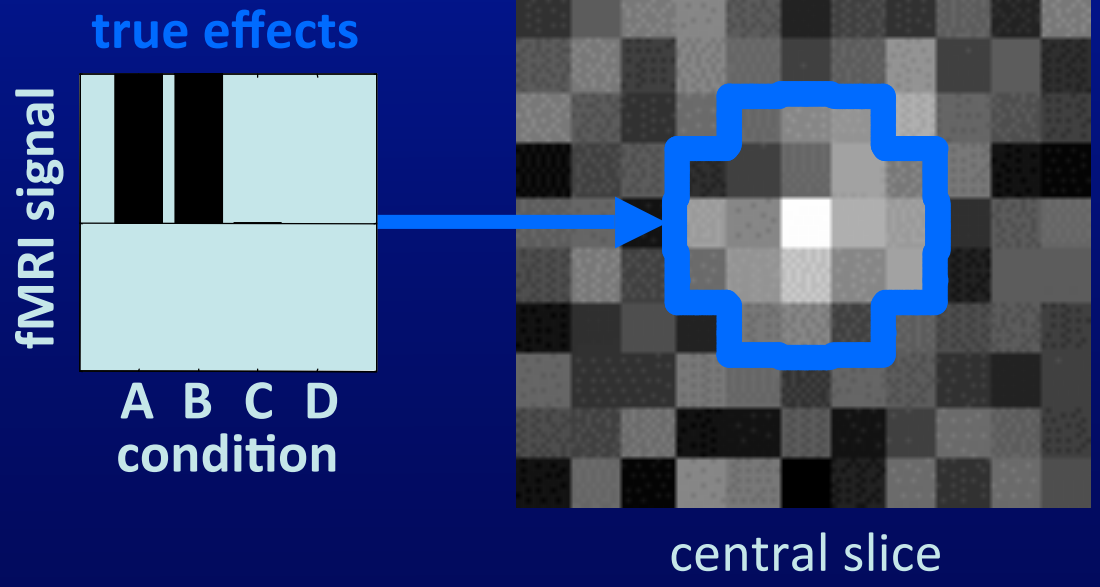
independent ROI

overfitted ROI

true region

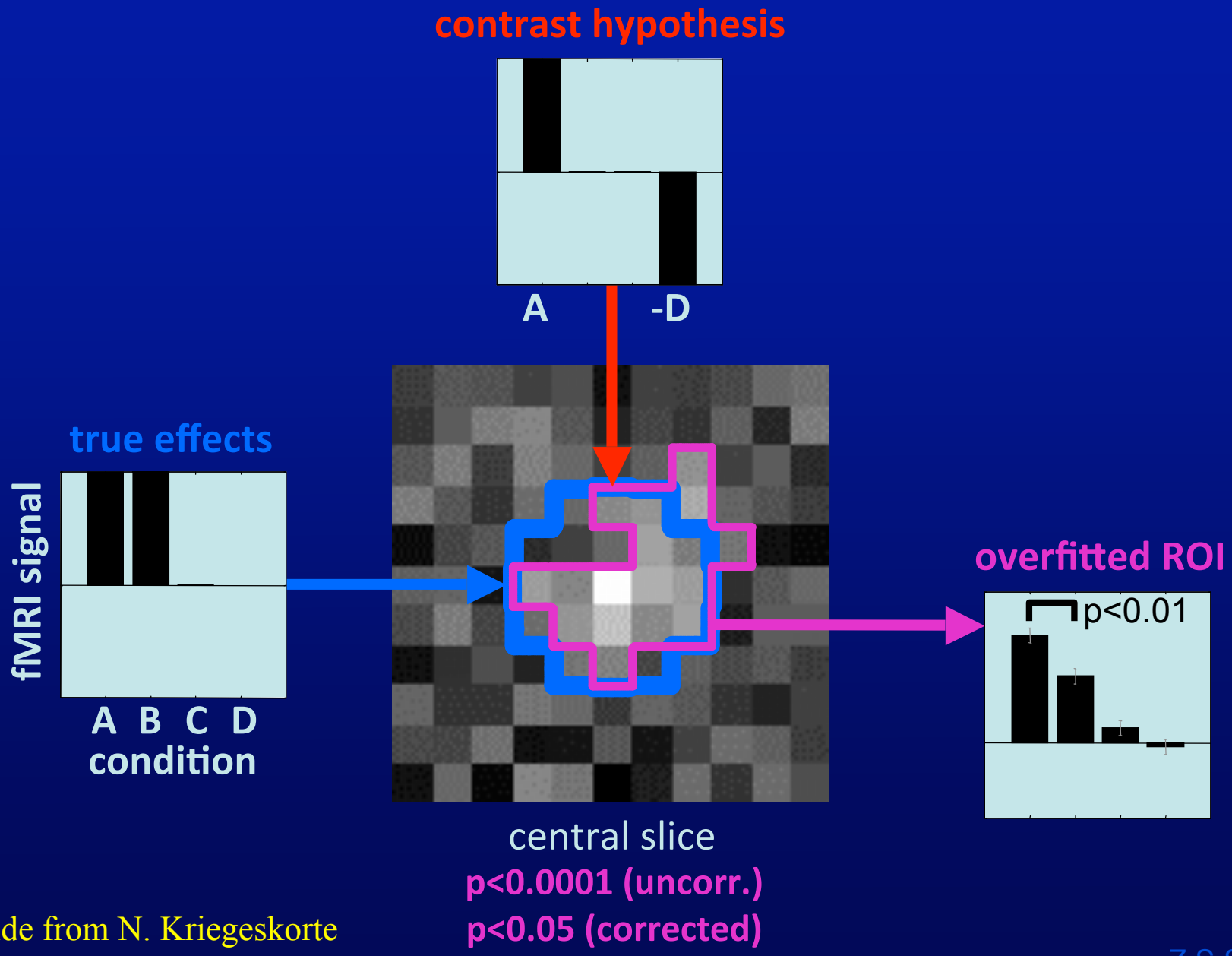
Slide from N. Kriegeskorte

Regional-average activation analysis



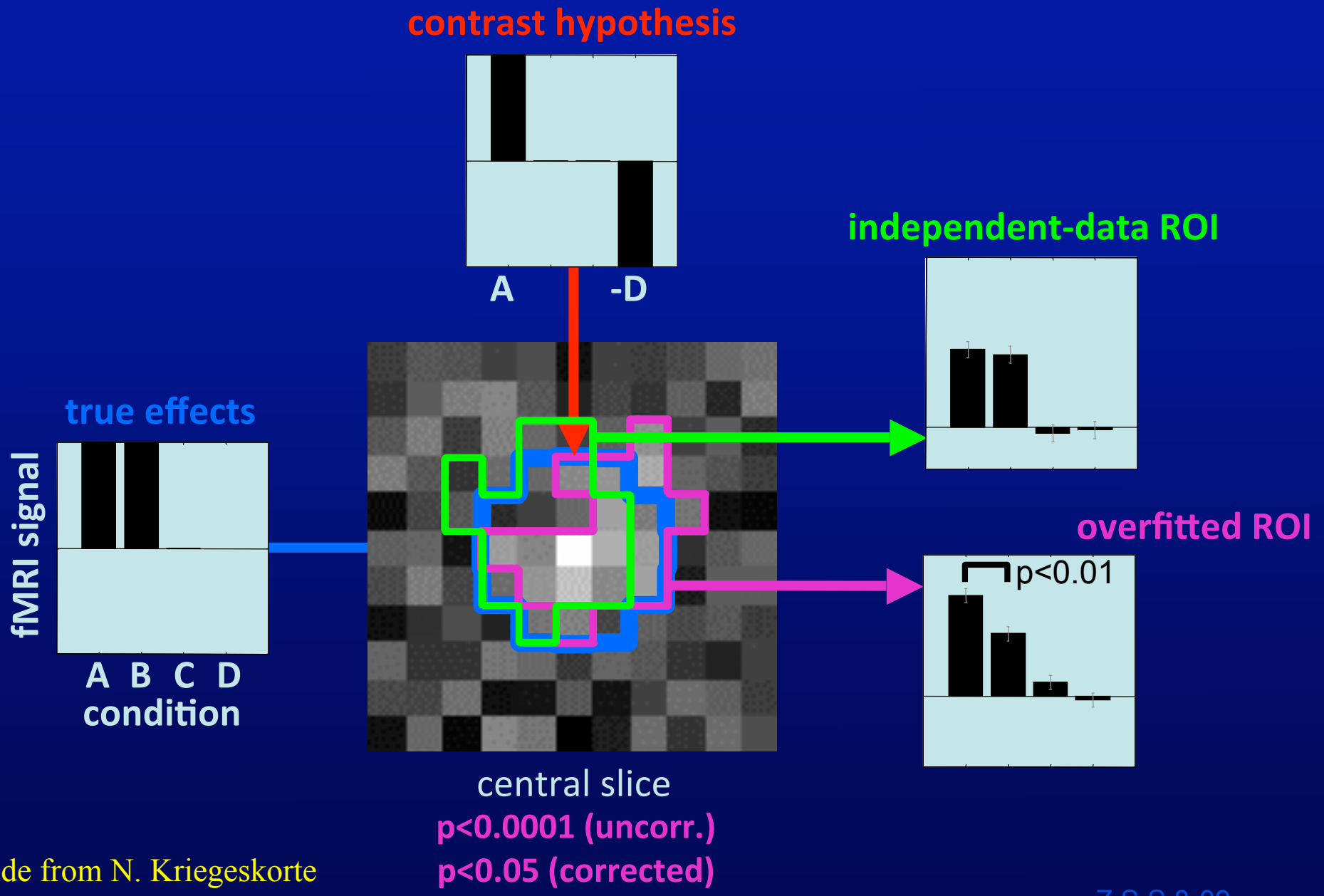
Slide from N. Kriegeskorte

Regional-average activation analysis



Slide from N. Kriegeskorte

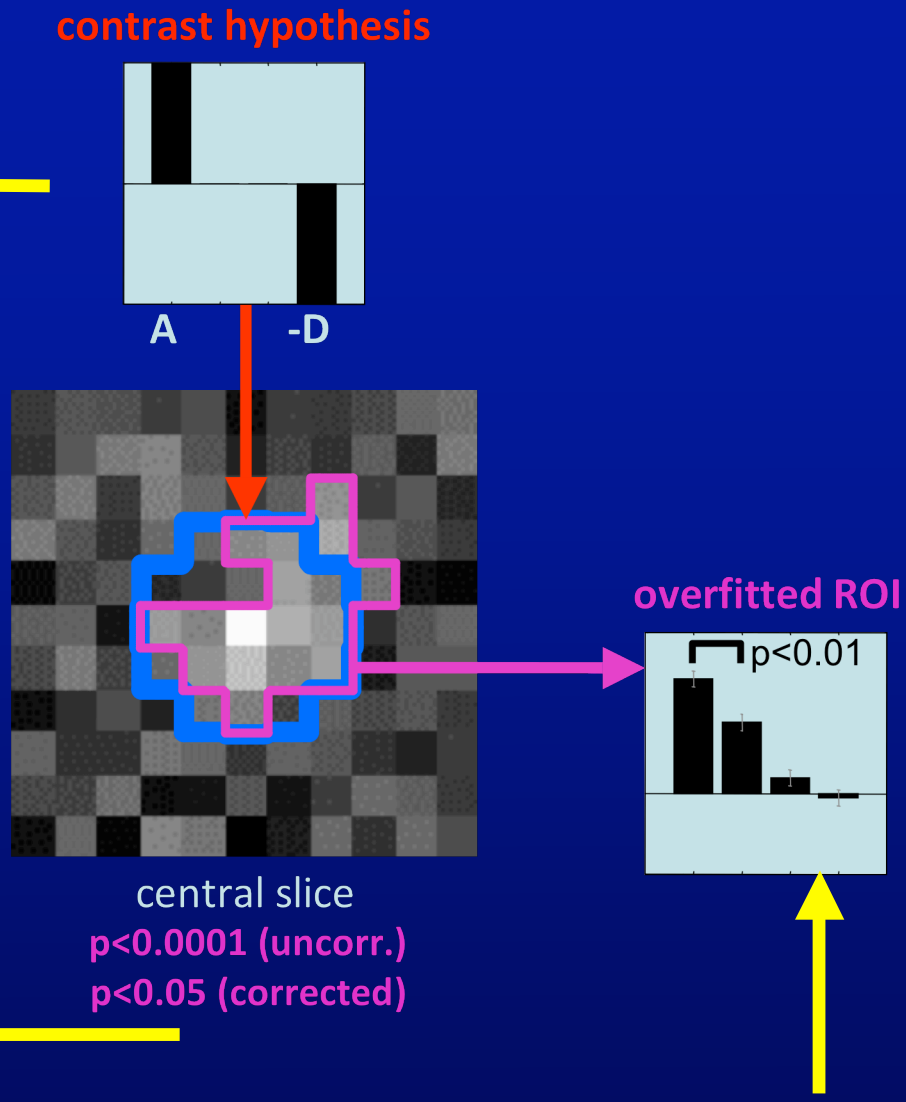
Regional-average activation analysis



Slide from N. Kriegeskorte

Regional-average activation analysis

(1)- Bias towards high values of A
AND
No bias towards B



(2)- Higher threshold (t) to meet multiple comparison correction
- The higher the threshold the worse the bias in (1)
(*problem made worse in low power cases*)

1 & 2 result in sample where A > B !

Notes on a scandal

- Beware selection bias / circularity
- Avoid hysteria
  - Circularity can lead to incorrect inferences
  - Ginormoulsy high correlations
    - Maybe caused by circularity
    - More likely caused by low power (very small number of subjects, many voxels)
    - But this does not mean that correlations do not exist!
      - But the average correlation of those that pass a high threshold can be much higher than true correlation
- Consider your false negatives

Notes on a scandal

- Beware selection bias / circularity
- Avoid hysteria
- Consider your false negatives
  - What's wrong with being 'safe'?
    - We are likely missing A LOT (>50%) of true positives
    - Wrong models about how the brain works !
      - Connectivity models very sensitive to the nodes in model
  - What can be done?
    - More power by better modeling signal AND noise
    - More power by having more subjects
    - Judicious covariate selection
    - Consider what happens at lower thresholds.
      - More nodes or just bigger blobs?

Acknowledgments

Robert Cox
Rick Reynolds
Gang Chen

Rasmus Birn
Niko Kriegeskorte
Marta Bianciardi
Der-Yow Chen
Kristina Simonyan
Michael Beauchamp

Julien Doyon